The Calcium-dependent protein kinase 1 from Toxoplasma gondii as target for structure-based drug design

Journal:	*Parasitology*
Manuscript ID	PAR-2017-0241.R1
Manuscript Type:	Special Issue Review (invited contributions only)
Date Submitted by the Author:	n/a
Complete List of Authors:	Cardew, Emily; Durham University Department of Chemistry, Chemistry Verlinde, Christophe; University of Washington School of Public Health, Biochemistry Pohl, Ehmke; Durham University Department of Chemistry, Chemistry
Key Words:	calcium-dependent protein kinase, CDPK1, Toxoplasma gondii, drug design, protein structure

Calcium-dependent protein kinases from *Toxoplasma gondii* as targets for structure-based drug design

Emily Cardew[1], Christophe L.M.J. Verlinde[2], Ehmke Pohl [1,3,4,*].

[1] Department of Biosciences, Durham University, Lower Mountjoy Durham DH1 3LE, UK

[2] Department of Biochemistry, University of Washington, Seattle, Washington, WA 98195, USA.

[3] Department of Chemistry, Durham University, South Road, Durham DH1, 3LE, UK.

[4] Biophysical Sciences Institute, Durham University, Durham DH1 3LE, UK.

* corresponding author, email: ehmke.pohl@durham.ac.uk

Summary

The apicomplexan protozoan parasites include the causative agents of animal and human diseases ranging from malaria (*Plasmodium* spp.) to toxoplasmosis (*Toxoplasma gondii*). The complex life cycle of *T. gondii* is regulated by a unique family of calcium-dependent protein kinases (CDPKs) that have become the target of intensive efforts to develop new therapeutics. In this review, we will summarize structure-based strategies and recent successes in the pursuit of specific and selective inhibitors of *T. gondii* CDPK1.

Introduction

The phylum of Apicomplexa contains approximately 6000 unicellular, eukaryotic parasites including *Plasmodium* spp, the causative agent of Malaria, and *Toxoplasma gondii*, responsible for toxoplasmosis in many important farm animals and humans (Sato, 2011). Morphologically, all members of the apicomplexan family share am distinctive apical complex and unique apical-localised organelles such as the apicoplast, is a non-photosynthetic relict plastid found in both, Plasmodium spp. and *T. gondii* (McFadden & Yeh, 2017). These parasites employ complex life cycles including both sexual and asexual reproduction, and often involving multiple hosts. *T. gondii*, first described in 1908 and often regarded as one of the most successful apicomplexan parasites, represents the key model organism of the phylum (Weiss & Dubey, 2009, Szabo & Finney, 2017, Dubey, 2008). Its primary hosts are members of the Felidae (cats) family and all other warm-blooded animals including humans, are intermediate hosts. It is estimated that up to one third of the human population is infected with *T. gondii* and thus are potential carriers. Although the infection is usually asymptotic in healthy individuals it can cause severe congenital disease during pregnancy (Kaye, 2011), and lead to life-threatening infections in immuno-compromised patients including those suffering from HIV, having received an organ transplant or receiving cancer chemotherapy treatment (Flegr *et al.*, 2014). Current toxoplasmosis treatment options are limited to a handful of antimicrobials such as sulphonamides, folic acid derivatives and certain macrolide antibiotics. However, these drugs often show limited efficacy and are associated with significant side effects (Alday & Doggett, 2017). Furthermore, there are no treatments available to target tissue cysts, the persistent form in which the parasite evades the host immune system, and to eradicate persistent *T. gondii* infections (Opsteegh *et al.*, 2015). Therefore, new drug targets and therapies are urgently needed. In addition to high-throughput screening approaches (Norcliffe *et al.*, 2014), structure-based methods in close combination with medicinal chemistry and biophysical and biological validation have become powerful tools in the search of new drugs against infectious diseases (Hol, 2015, Groftehauge *et al.*, 2015, Muller, 2017, Verlinde *et al.*, 2009).

The role of Calcium-dependent protein kinases

Calcium is an essential element for almost all eukaryotic organisms with wide-ranging biological functions. In *Toxoplasma,* Ca^{2+}-ions play a key roles in cell signalling and in

pathogen-host interaction including cell invasion, motility of the parasite within the host and differentiation during the parasites complex life cycle (Irvine, 1986, Nagamune *et al.*, 2008, Lourido & Moreno, 2015). Calcium dependent protein kinases (CDPKs) are serine/threonine kinases that are only found in plants and protists including ciliates and apicomplexan parasites. Importantly, they provide the mechanistic link between calcium signalling and motility, differentiation and invasion (Tzen *et al.*, 2007, Billker *et al.*, 2009). These key roles of CDPKs have been proven in a range of knock-out studies in various species and underline their potential as targets for novel therapeutics (Long *et al.*, 2016) (Wang *et al.*, 2016). So far, at least twelve different CDPKs have been putatively identified in the *T. gondii* alone ranging from 583 (CDPK1) to more than 2000 (CDPK7, CDPK8 (Morlon-Guyot *et al.*, 2014)) amino acids in length with sequence identities ranging from 51% (CDPK1 and CDPK3 (Treeck *et al.*, 2014)) to lower than 10% in the conserved regions (Table 1) (Hui *et al.*, 2015). CDPKs are members of the Calmodulin/Calcium kinase (CaM) family and hence they share an N-terminal kinase domain (KD) linked via a junctional domain to a series of C-terminal Calcium-binding motifs. However, as evidenced by their sequence variation, different members of the CDPK family have vastly different substrates and biological functions in *T. gondii* biology. CDPK1 which is the most comprehensively studied member of the family, has been shown to be required for the microneme secretion at the apical complex and parasite proliferation (Lourido *et al.*, 2010, Child *et al.*, 2017).

Due to its key role in infection and because the mammalian hosts do not possess any representative of the same kinase family, CDPK1 from Plasmodium Cryptosporidium and Toxoplasma spp. has attracted significant attention as a potential novel drug target (Donald *et al.*, 2006, Sugi *et al.*, 2010, Larson *et al.*, 2012). Here we will review strategies and recent results in the discovery, design and potency of inhibitors of the CDPK1 from *T. gondii* (*Tg*CDPK1).

Activation of *Tg*CDPK1 by Calcium

Detailed structural studies began in 2010 when the crystal structures of both the auto-inhibited and the Ca^{2+}-activated forms of *Tg*CDPK1 were published (Ojo *et al.*, 2010, Wernimont *et al.*, 2010). These structures revealed the expected canonical KD in similar overall conformations, however, the Ca^{2+}-binding domain (also designated CPDK activating domain or CAD) adopted two vastly different conformations and orientations (Figure 1a and

1b). In its inactive state the CAD (shown in raspberry red) adopts an elongated form reminiscent of apo-calmodulin starting with a long helix followed by the first Ca^{2+}-binding motif (EF-hands) which is connected via another long helix to the second C-terminal EF-hand. The first long helix is responsible for the auto-inhibitory effect by blocking the substrate binding site and providing a basic lysine residue to bind a cluster of conserved acidic residues. Calcium binding leads to a dramatic rearrangement and refolding of the protein chain (Figure 1b) (Wernimont *et al.*, 2010). The entire regulatory domain is shifted to the other side of the protein hence liberating the active site of the kinase domain as shown in Figure 2. In addition, the regulatory calcium-binding domain is collapsed so that the two long helices are no longer arranged in an anti-parallel fashion but are partially unwound and interwoven to form a more globular overall shape. These structural changes are reminiscent to the calcium-bound structure of calmodulin (Kursula, 2014). However, the reorientation and structural changes observed in *Tg*CDPK1 are more profound, presumably due to the long linker region between the two Ca^{2+}-binding EF-hands.

Comparison with human kinases

Historically characterising (protozoan) kinases as potential drug targets and developing selective inhibitors has been considered challenging due to the fact that the overall protein fold and the active sites are structurally well conserved in all kinases. The structural similarities are obvious when comparing the crystal structures of the kinase domain of CDPK1 *from T. gondii* with Calcium/Calmodulin (CaM) dependent-kinase II from *H. sapiens* (*Hs*CaMKII) (Figure 3a). These two proteins, which share a sequence identity of approximately 42% over 264 residues of the kinase domain display the same canonical kinase fold and superimpose with an overall root mean square deviation of approximately 1.5 Å. Note that the loop over the adenosine triphosphate (ATP) binding site adopts a very different conformation presumably due to an induced fit of binding of two very different ligands. *Tg*CDPK1 is bound to the ATP analogue ANP while *Hs*CaMKII is bound to a comparatively small inhibitor. More importantly there are significant differences in the ATP binding site, specifically an unusually small residue (glycine) close to the adenine binding position. This residue, glycine 128 is also termed the *gatekeeper* residue. Almost all mammalian kinases possess a large residue, a phenylalanine in *Hs*CaMKII for example, in this position. Hence, the protozoan kinases feature an enlarged ATP binding site with a

hydrophobic pocket that can be exploited for structure-based drug design. This key structural difference in the binding pocket is shown in the surface representation where the ATP-analogue is shown as stick representation (Figure 3a). The additional space at the end of the pocket below the surface of the gatekeeper residue glycine 128 in magenta is clearly visible.

Development of specific *Tg*CDPK1 inhibitors

Soon after the importance of this enzyme and the structural differences were established two groups started to develop selective *Tg*CDPK1 inhibitors. Initial compounds were based on known inhibitors previously developed for yeast kinases featuring amino acids with small side chains at the *gatekeeper* position. Importantly, these known kinase inhibitors, termed *bumped kinase inhibitors* (BKI) have been shown to be inactive against mammalian kinases (Hanke *et al.*, 1996). Generally, BKIs are based on the planar pyrazolo[3,4-d]pyrimidin-4-amine substituted with a bulky hydrophobic group on the C3 position (Bishop *et al.*, 1998). The first example of a BKI with a sub-micromolar IC_{50} is 1-(1-methylethyl)-3-(naphthalen-1-ylmethyl)-1H-pyrazolo[3,4-d] pyrimidin-4-amine. The co-crystal structure shows that the naphtalen-1-ylmethyl- moiety perfectly fills the hydrophobic pocket created by the small gatekeeper residue Gly128 and lined by methionine and leucine residues, and one lysine residue (Figure 4a,c). The chemically closely related 1-tert-butyl-3-naphthalen-2-yl-1H-pyrazolo[3,4-d]pyrimidin-4-amine (Figure 4b,d) adopts a similar conformation with the bulky aromatic substituent at the C3 position occupying the space next to the gatekeeper residue. Critically for the subsequent drug development was the fact that these and related BKIs reduced *T. gondii* proliferation significantly (Ojo *et al.*, 2010, Sugi *et al.*, 2010). These results sparked extensive medicinal chemistry efforts where a large number of compounds based on the BKI scaffold (4-amino-1H-pyrazole[3,4-d]pyrimidine) were synthesized and tested resulting in optimized *Tg*CDPK1 inhibitors. A number of compounds exhibited sub- or low-nanomolar for IC_{50} values and high activity in parasite growth models (EC_{50} in the low- and sub micromolar range) while retaining specificity when compared to mammalian kinases (Lourido *et al.*, 2013) (Zhang *et al.*, 2014) (Moine *et al.*, 2015). In addition to the pyrazolopyrimidine (PP) scaffolds, acylbenzimidazole and 5-aminopyrzazole-4-carboxamide based compounds shown in Figure 5 with similar properties have been successfully developed (Zhang *et al.*, 2012, Zhang *et al.*, 2014, Huang *et al.*, 2015). While the initial BKIs

showed excellent potency *in vitro* and *in vivo* they also exhibited significant hERG (**h**uman **E**ther-a go-go-go **R**elated **G**ene) inhibition thus posing potential cardiotoxicity (Doggett *et al.*, 2014). Further extensive medicinal chemistry efforts finally led to the current lead of *Tg*CDPK1 inhibitors, (1-{4-amino-3-[2-(cyclopropyloxy)quinolin-6-yl]-1H-pyrazolo[3,4-d]pyrimidin-1-yl}-2-methylpropan-2-ol) that combined high activity and selectivity with favourable pharmacokinetic properties and low hERG activity (Vidadala *et al.*, 2016). Note that the copounds is bound to the protein via H-bonds of the pyrimidin ring to the main chain of the proitein, while the hydrophobic cyclopropyloxyquinolin moiety forms a large number of hydrophobic interactions. Taken together, the structure based approaches of drug development applied to *Tg*CDPK1 has led to three different series of compounds with high inhibitory activity, good pharmacokinetic parameters and promising efficacy in murine models.

Future challenges

Over the last five years there has been significant progress in the development of selective inhibitors of one of the key CDPKs from *T. gondii* taking advantage of a series of high-resolution crystal structures. Although the most promising compounds show high efficacy in murine models more work needs to be done to increase solubility and bio-availability in order to proceed to clinical trials. While most of the previous work has focused on *T. gondii*, further work is currently underway to investigate inhibitors of CDPK1 from Cryptosporidium and Plasmodium spp. (Gaji *et al.*, 2014, Green *et al.*, 2015, Crowther *et al.*, 2016). In addition, more works needs to be done to understand the role of the other members of the Apicomplexan CDPK family. In this regard, the development of CRISPR/Cas9 technology in members of Apicomplexan family (Shen *et al.*, 2014, Vinayak *et al.*, 2015) facilitated the detailed analysis of the biological function of CDPK family members (Long *et al.*, 2016).

Acknowledgments

This work was generously supported by the BBSRC grant BB/M024156/1. We would like to thank the Biophysical Sciences Institute (BSI) for seed corn funding and the Advanced Studies (IAS) for a Senior Research Fellowship at Durham University to CLMV. EMC is grateful for a studentship by the Newcastle-Liverpool-Durham BBSRC Doctoral Training Partnership. We would like to thank R. Eno and I. Edwards for excellent technical support and many fruitful discussions.

References

Alday, P. H. & Doggett, J. S. (2017). *Drug Des Devel Ther* **11**, 273-293.

Billker, O., Lourido, S. & Sibley, L. D. (2009). *Cell Host Microbe* **5**, 612-622.

Bishop, A. C., Shah, K., Liu, Y., Witucki, L., Kung, C. & Shokat, K. M. (1998). *Curr Biol* **8**, 257-266.

Child, M. A., Garland, M., Foe, I., Madzelan, P., Treeck, M., van der Linden, W. A., Oresic Bender, K., Weerapana, E., Wilson, M. A., Boothroyd, J. C., Reese, M. L. & Bogyo, M. (2017). *MBio* **8**.

Crowther, G. J., Hillesland, H. K., Keyloun, K. R., Reid, M. C., Lafuente-Monasterio, M. J., Ghidelli-Disse, S., Leonard, S. E., He, P., Jones, J. C., Krahn, M. M., Mo, J. S., Dasari, K. S., Fox, A. M., Boesche, M., El Bakkouri, M., Rivas, K. L., Leroy, D., Hui, R., Drewes, G., Maly, D. J., Van Voorhis, W. C. & Ojo, K. K. (2016). *PLoS One* **11**, e0149996.

Doggett, J. S., Ojo, K. K., Fan, E., Maly, D. J. & Van Voorhis, W. C. (2014). *Antimicrob Agents Chemother* **58**, 3547-3549.

Donald, R. G., Zhong, T., Wiersma, H., Nare, B., Yao, D., Lee, A., Allocco, J. & Liberator, P. A. (2006). *Mol Biochem Parasitol* **149**, 86-98.

Dubey, J. P. (2008). *J Eukaryot Microbiol* **55**, 467-475.

Flegr, J., Prandota, J., Sovickova, M. & Israili, Z. H. (2014). *PLoS One* **9**, e90203.

Gaji, R. Y., Checkley, L., Reese, M. L., Ferdig, M. T. & Arrizabalaga, G. (2014). *Antimicrob Agents Chemother* **58**, 2598-2607.

Green, J. L., Moon, R. W., Whalley, D., Bowyer, P. W., Wallace, C., Rochani, A., Nageshan, R. K., Howell, S. A., Grainger, M., Jones, H. M., Ansell, K. H., Chapman, T. M., Taylor, D. L., Osborne, S. A., Baker, D. A., Tatu, U. & Holder, A. A. (2015). *Antimicrob Agents Chemother* **60**, 1464-1475.

Groftehauge, M. K., Hajizadeh, N. R., Swann, M. J. & Pohl, E. (2015). *Acta Crystallogr D Biol Crystallogr* **71**, 36-44.

Hanke, J. H., Gardner, J. P., Dow, R. L., Changelian, P. S., Brissette, W. H., Weringer, E. J., Pollok, B. A. & Connelly, P. A. (1996). *J Biol Chem* **271**, 695-701.

Hol, W. G. (2015). *Acta Crystallogr F Struct Biol Commun* **71**, 485-499.

Huang, W., Ojo, K. K., Zhang, Z., Rivas, K., Vidadala, R. S., Scheele, S., DeRocher, A. E., Choi, R., Hulverson, M. A., Barrett, L. K., Bruzual, I., Siddaramaiah, L. K., Kerchner, K. M., Kurnick, M. D., Freiberg, G. M., Kempf, D., Hol, W. G., Merritt, E. A., Neckermann, G., de Hostos, E. L., Isoherranen, N., Maly, D. J., Parsons, M., Doggett, J. S., Van Voorhis, W. C. & Fan, E. (2015). *ACS Med Chem Lett* **6**, 1184-1189.

Hui, R., El Bakkouri, M. & Sibley, L. D. (2015). *Trends Pharmacol Sci* **36**, 452-460.

Irvine, R. F. (1986). *Br Med Bull* **42**, 369-374.

Kaye, A. (2011). *J Pediatr Health Care* **25**, 355-364.

Kursula, P. (2014). *Amino Acids* **46**, 2295-2304.

Larson, E. T., Ojo, K. K., Murphy, R. C., Johnson, S. M., Zhang, Z., Kim, J. E., Leibly, D. J., Fox, A. M., Reid, M. C., Dale, E. J., Perera, B. G., Kim, J., Hewitt, S. N., Hol, W. G., Verlinde, C. L., Fan, E., Van Voorhis, W. C., Maly, D. J. & Merritt, E. A. (2012). *J Med Chem* **55**, 2803-2810.

Long, S., Wang, Q. & Sibley, L. D. (2016). *Infect Immun* **84**, 1262-1273.

Lourido, S., Jeschke, G. R., Turk, B.E. & Sibley, D. (2013). *ACS Chemical Biology* **8**, 1155-1162.

Lourido, S. & Moreno, S. N. (2015). *Cell Calcium* **57**, 186-193.

Lourido, S., Shuman, J., Zhang, C., Shokat, K. M., Hui, R. & Sibley, L. D. (2010). *Nature* **465**, 359-362.
McFadden, G. I. & Yeh, E. (2017). *Int J Parasitol* **47**, 137-144.
Moine, E., Dimier-Poisson, I., Enguehard-Gueiffier, C., Loge, C., Penichon, M., Moire, N., Delehouze, C., Foll-Josselin, B., Ruchaud, S., Bach, S., Gueiffier, A., Debierre-Grockiego, F. & Denevault-Sabourin, C. (2015). *Eur J Med Chem* **105**, 80-105.
Morlon-Guyot, J., Berry, L., Chen, C. T., Gubbels, M. J., Lebrun, M. & Daher, W. (2014). *Cell Microbiol* **16**, 95-114.
Muller, I. (2017). *Acta Crystallogr D Struct Biol* **73**, 79-92.
Nagamune, K., Moreno, S. N., Chini, E. N. & Sibley, L. D. (2008). *Subcell Biochem* **47**, 70-81.
Norcliffe, J. L., Alvarez-Ruiz, E., Martin-Plaza, J. J., Steel, P. G. & Denny, P. W. (2014). *Parasitology* **141**, 8-16.
Ojo, K. K., Larson, E. T., Keyloun, K. R., Castaneda, L. J., Derocher, A. E., Inampudi, K. K., Kim, J. E., Arakaki, T. L., Murphy, R. C., Zhang, L., Napuli, A. J., Maly, D. J., Verlinde, C. L., Buckner, F. S., Parsons, M., Hol, W. G., Merritt, E. A. & Van Voorhis, W. C. (2010). *Nat Struct Mol Biol* **17**, 602-607.
Opsteegh, M., Kortbeek, T. M., Havelaar, A. H. & van der Giessen, J. W. (2015). *Clin Infect Dis* **60**, 101-107.
Rellos, P., Pike, A. C., Niesen, F. H., Salah, E., Lee, W. H., von Delft, F. & Knapp, S. (2010). *PLoS Biol* **8**, e1000426.
Sato, S. (2011). *Cell Mol Life Sci* **68**, 1285-1296.
Shen, B., Brown, K. M., Lee, T. D. & Sibley, L. D. (2014). *MBio* **5**, e01114-01114.
Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J. D. & Higgins, D. G. (2011). *Mol Syst Biol* **7**, 539.
Sugi, T., Kato, K., Kobayashi, K., Watanabe, S., Kurokawa, H., Gong, H., Pandey, K., Takemae, H. & Akashi, H. (2010). *Eukaryot Cell* **9**, 667-670.
Szabo, E. K. & Finney, C. A. (2017). *Trends Parasitol* **33**, 113-127.
Treeck, M., Sanders, J. L., Gaji, R. Y., LaFavers, K. A., Child, M. A., Arrizabalaga, G., Elias, J. E. & Boothroyd, J. C. (2014). *PLoS Pathog* **10**, e1004197.
Tzen, M., Benarous, R., Dupouy-Camet, J. & Roisin, M. P. (2007). *Parasite* **14**, 141-147.
Verlinde, C. L., Fan, E., Shibata, S., Zhang, Z., Sun, Z., Deng, W., Ross, J., Kim, J., Xiao, L., Arakaki, T. L., Bosch, J., Caruthers, J. M., Larson, E. T., Letrong, I., Napuli, A., Kelly, A., Mueller, N., Zucker, F., Van Voorhis, W. C., Buckner, F. S., Merritt, E. A. & Hol, W. G. (2009). *Curr Top Med Chem* **9**, 1678-1687.
Vidadala, R. S., Rivas, K. L., Ojo, K. K., Hulverson, M. A., Zambriski, J. A., Bruzual, I., Schultz, T. L., Huang, W., Zhang, Z., Scheele, S., DeRocher, A. E., Choi, R., Barrett, L. K., Siddaramaiah, L. K., Hol, W. G., Fan, E., Merritt, E. A., Parsons, M., Freiberg, G., Marsh, K., Kempf, D. J., Carruthers, V. B., Isoherranen, N., Doggett, J. S., Van Voorhis, W. C. & Maly, D. J. (2016). *J Med Chem* **59**, 6531-6546.
Vinayak, S., Pawlowic, M. C., Sateriale, A., Brooks, C. F., Studstill, C. J., Bar-Peled, Y., Cipriano, M. J. & Striepen, B. (2015). *Nature* **523**, 477-480.
Wang, J. L., Huang, S. Y., Li, T. T., Chen, K., Ning, H. R. & Zhu, X. Q. (2016). *Parasitol Res* **115**, 697-702.
Weiss, L. M. & Dubey, J. P. (2009). *Int J Parasitol* **39**, 895-901.

Wernimont, A. K., Artz, J. D., Finerty, P., Jr., Lin, Y. H., Amani, M., Allali-Hassani, A., Senisterra, G., Vedadi, M., Tempel, W., Mackenzie, F., Chau, I., Lourido, S., Sibley, L. D. & Hui, R. (2010). *Nat Struct Mol Biol* **17**, 596-601.

Zhang, Z., Ojo, K. K., Johnson, S. M., Larson, E. T., He, P., Geiger, J. A., Castellanos-Gonzalez, A., White, A. C., Jr., Parsons, M., Merritt, E. A., Maly, D. J., Verlinde, C. L., Van Voorhis, W. C. & Fan, E. (2012). *Bioorg Med Chem Lett* **22**, 5264-5267.

Zhang, Z., Ojo, K. K., Vidadala, R., Huang, W., Geiger, J. A., Scheele, S., Choi, R., Reid, M. C., Keyloun, K. R., Rivas, K., Siddaramaiah, L. K., Comess, K. M., Robinson, K. P., Merta, P. J., Kifle, L., Hol, W. G., Parsons, M., Merritt, E. A., Maly, D. J., Verlinde, C. L., Van Voorhis, W. C. & Fan, E. (2014). *ACS Med Chem Lett* **5**, 40-44.

Table 1: The protein sequence identities between the twelve full length putati8ve CDPKs of *T. gondii*, rounded to the nearest whole number, derived from a multiple sequence alignment (MSA) generated using Clustal Omega (Sievers *et al.*, 2011).

	CDPK1	CDPK2	CDPK2A	CDPK2B	CDPK3	CDPK4	CDPK4A	CDPK5	CDPK6	CDPK7	CDPK8	CDPK9
CDPK1		22%	23%	30%	51%	14%	24%	25%	9%	7%	6%	22%
CDPK2	22%		34%	34%	25%	14%	21%	32%	14%	7%	9%	26%
CDPK2A	23%	34%		37%	27%	20%	20%	32%	15%	9%	9%	24%
CDPK2B	30%	34%	37%		33%	16%	25%	36%	13%	8%	8%	26%
CDPK3	51%	25%	27%	33%		14%	25%	30%	12%	7%	7%	25%
CDPK4	14%	14%	20%	16%	14%		14%	15%	17%	11%	9%	14%
CDPK4A	24%	21%	20%	25%	25%	14%		21%	8%	6%	6%	18%
CDPK5	25%	32%	32%	36%	30%	15%	21%		13%	7%	8%	26%
CDPK6	9%	14%	15%	13%	12%	17%	8%	13%		16%	8%	12%
CDPK7	7%	7%	9%	8%	7%	11%	6%	7%	16%		8%	7%
CDPK8	6%	9%	9%	8%	7%	9%	6%	8%	8%	8%		7%
CDPK9	22%	26%	24%	26%	25%	14%	18%	26%	12%	7%	7%	

Figure Legends

Figure 1: Ribbon representation of the crystal structure of CDPK1 from *T. gondii* with the kinase domain depicted in cyan, the regulatory domain in raspberry red and the non-hydrolysable ligand ANP in stick representation (a) CDPK1 in its inactive auto-inhibited state (PDB code: 3KU2) (Wernimont *et al.*, 2010) (b) CDPK1 in its calcium-bound activated state with the Ca^{2+}-ions shown as green spheres (PDB code: 3HX4) (Wernimont *et al.*, 2010).

Figure 2: Ribbon diagram of the least-squares superposition of the inactive and active forms of CDPK1 with the kinase domains shown in different shades of cyan, the regulatory domain in shades of red, respectively. Only the kinase domain was used to calculate the transformation matrix which was then applied to the entire protein chain.

Figure 3: (a) Least squares superposition of the kinase domain of *Tg*CDPK1 (depicted in cyan) in its active form (PDB: 3HX4) with *Hs*CaMKII bound to an inhibitor (PDB: 2VZ6) (shown in green) (Rellos *et al.*, 2010). The non-hydrolysable ATP analogue bound in CDPK1 is presented as ball-and-stick representation to highlight the ATP binding site. (b) Surface representation of *Tg*CDPK1 viewing into the binding pocket with color coding according to atom type (oxygen in red, nitrogen in blue, carbon in grey). The surface of Gly128 (gatekeeper residue) is shown in magenta highlighting the additional space in the binding pocket of *Tg*CDPK1.

Figure 4: Close-up of BKIs bound to *Tg*CDPK1 in the ATP binding site. The gatekeeper residue Gly128 is depicted in magenta, key hydrophobic residue of the binding site in grey (a) 1-(1-methylethyl)-3-(naphthalen-1-ylmethyl)-1H-pyrazolo[3,4-d]pyrimidin-4-amine shown in ball-and-stick representation (PDB: 3i7b) (b) 1-tert-butyl-3-naphthalen-2-yl-1H-pyrazolo[3,4-d]pyrimidin-4-amine (PDB:3i7c) (Ojo *et al.*, 2010). (c) (d) chemical structures of the ligands.

Figure 5: The three different scaffolds for *Tg*CDPK1 inhibitors (a) Pyrazolpyrimidines (b) Acylbenzimidazoles (c) 5-aminopyrazole-4-carboxamide

Figure 6: Crystal structure of (1-{4-amino-3-[2-(cyclopropyloxy)quinolin-6-yl]-1H-pyrazolo[3,4-d]pyrimidin-1-yl}-2-methylpropan-2-ol) shown in stick representation bound to for *Tg*CDPK1 shown in cartoon representation with selected residues depicted in sticks (Vidadala *et al.*, 2016).

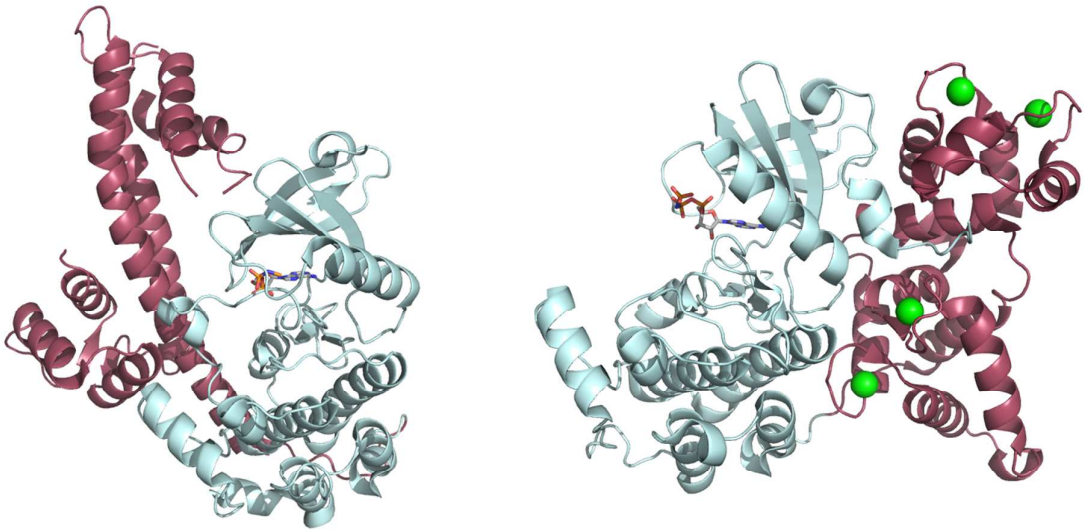
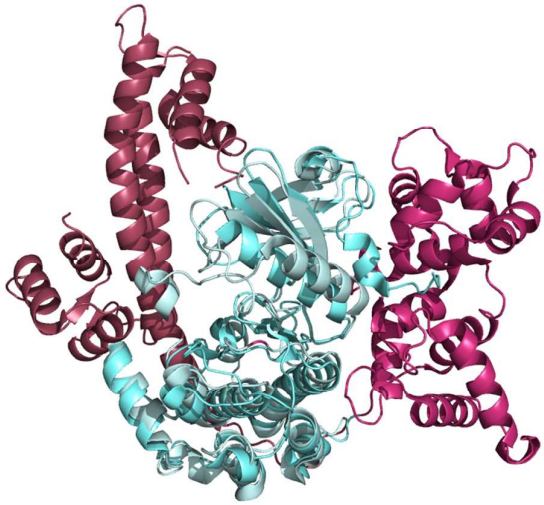


Figure 1

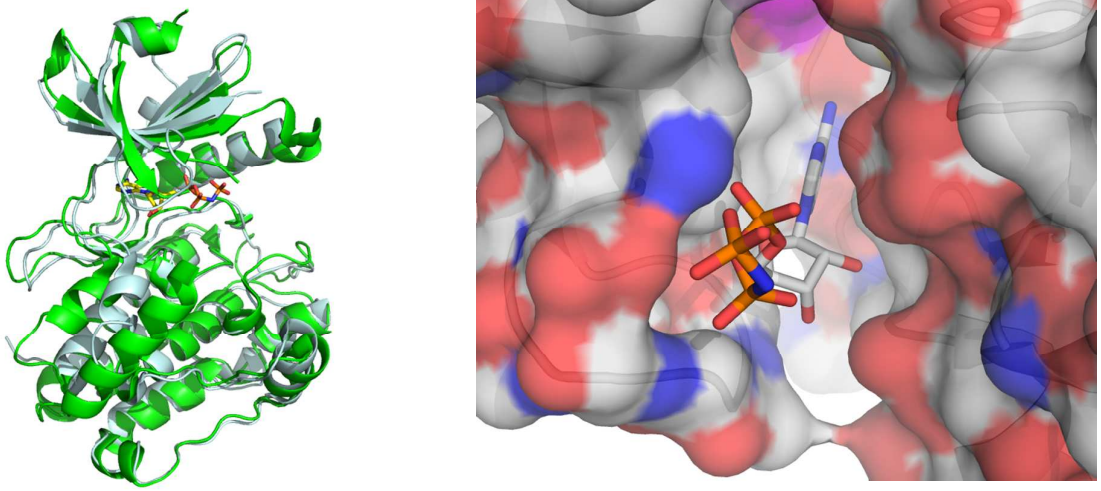
(a) (b)

Figure 2



339 **Figure 3**

340 (a) (b)

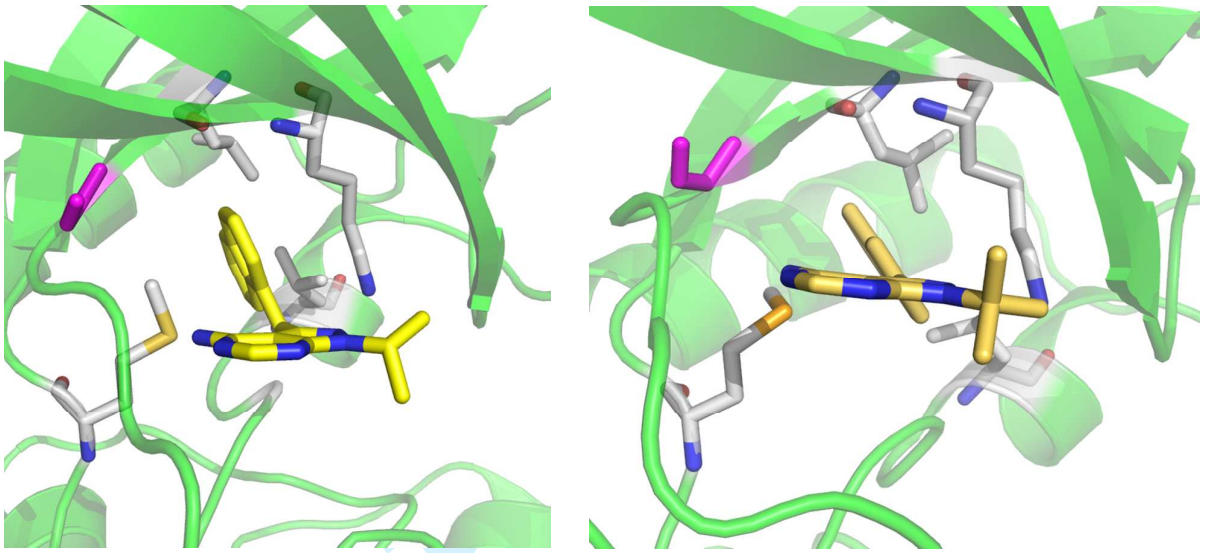


341

342

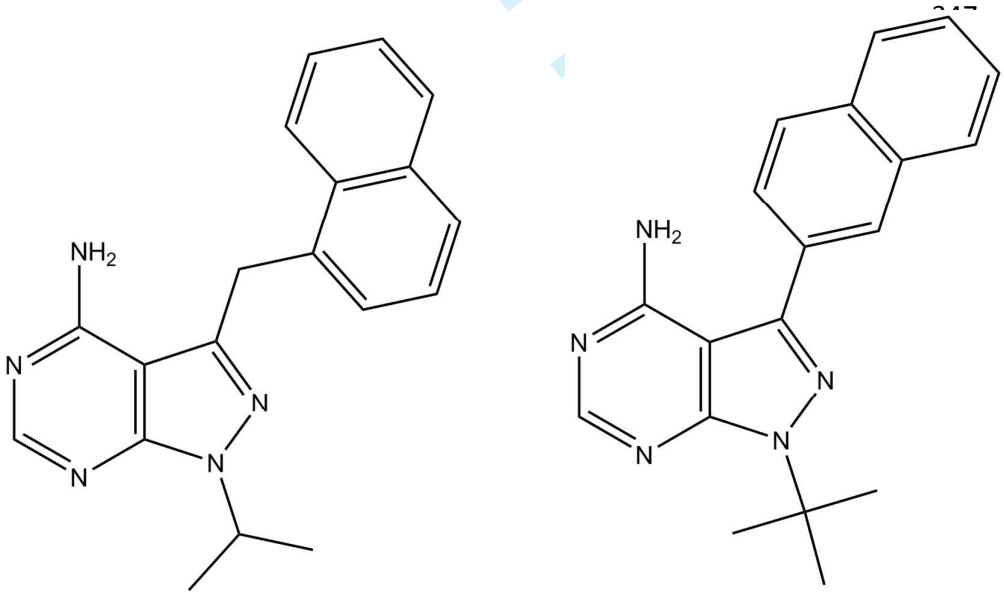
343 **Figure 4**

344 (a) (b)



345

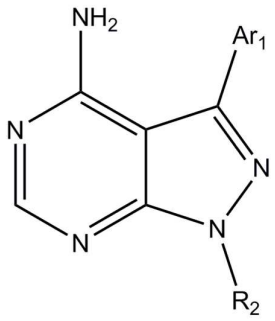
346 (c) (d)



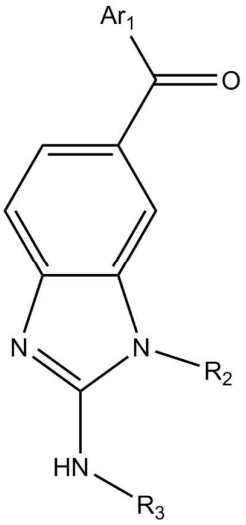
351
352
353

Figure 5

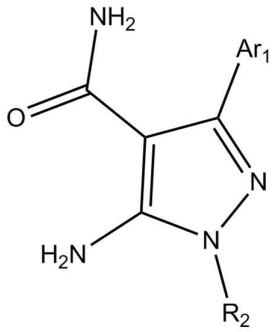
(a) Pyrazolpyrimidines



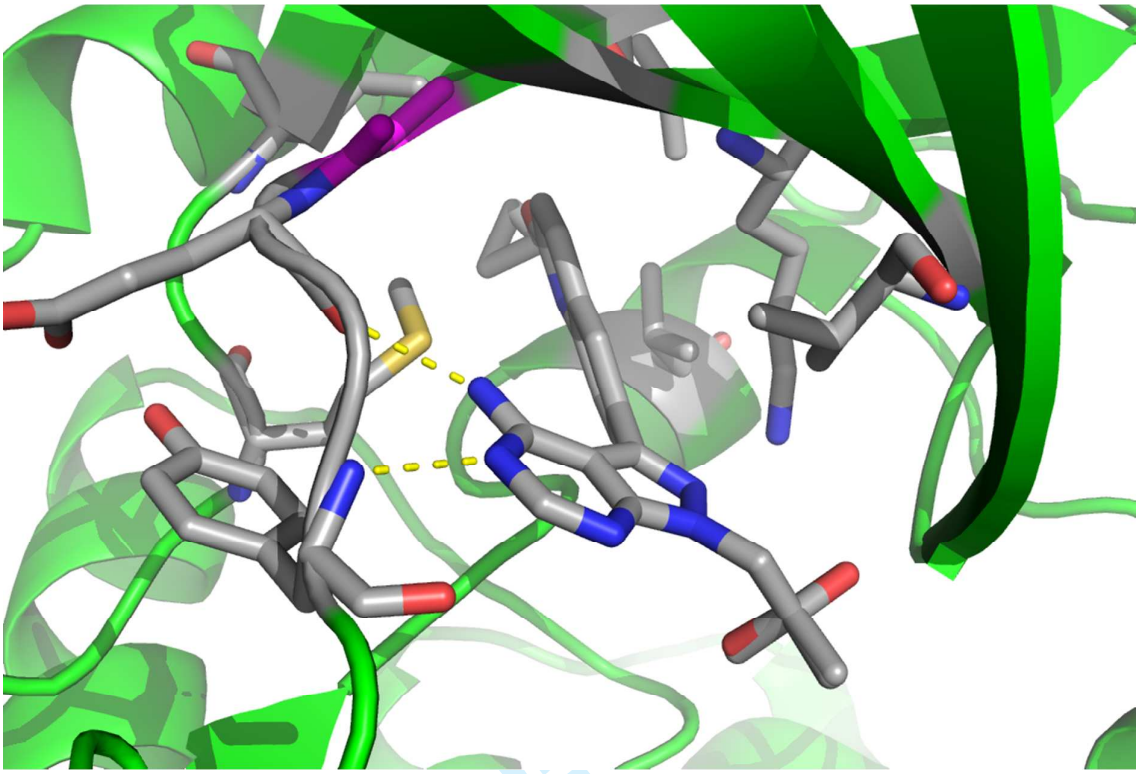
(b) Acylbenzimidazoles



(c) 5-aminopyrazole-4-carboxamide



375 Figure 6.



376

377

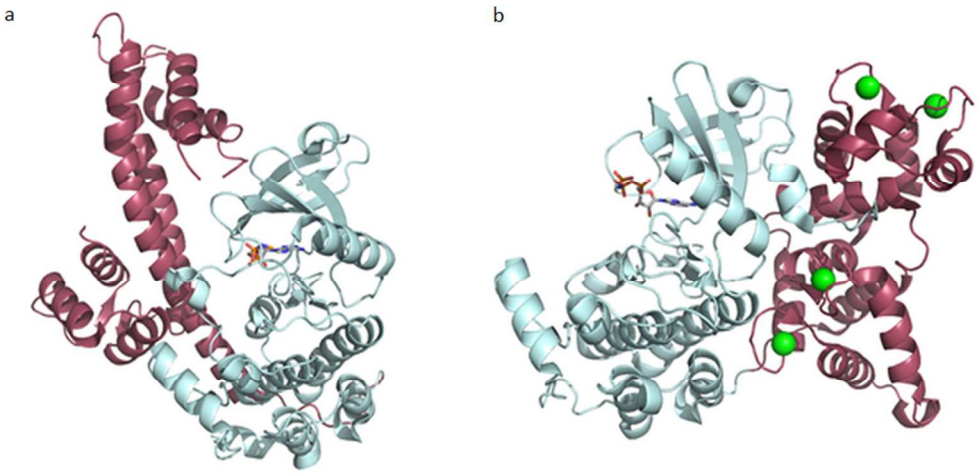


Figure 1: Ribbon representation of the crystal structure of CDPK1 from T. gondii with the kinase domain depicted in cyan, the regulatory domain in raspberry red and the non-hydrolysable ligand ANP in stick representation (a) CDPK1 in its inactive auto-inhibited state (PDB code: 3KU2) (Wernimont et al., 2010) (b) CDPK1 in its calcium-bound activated state with the Ca2+-ions shown as green spheres (PDB code: 3HX4) (Wernimont et al., 2010).

78x38mm (300 x 300 DPI)

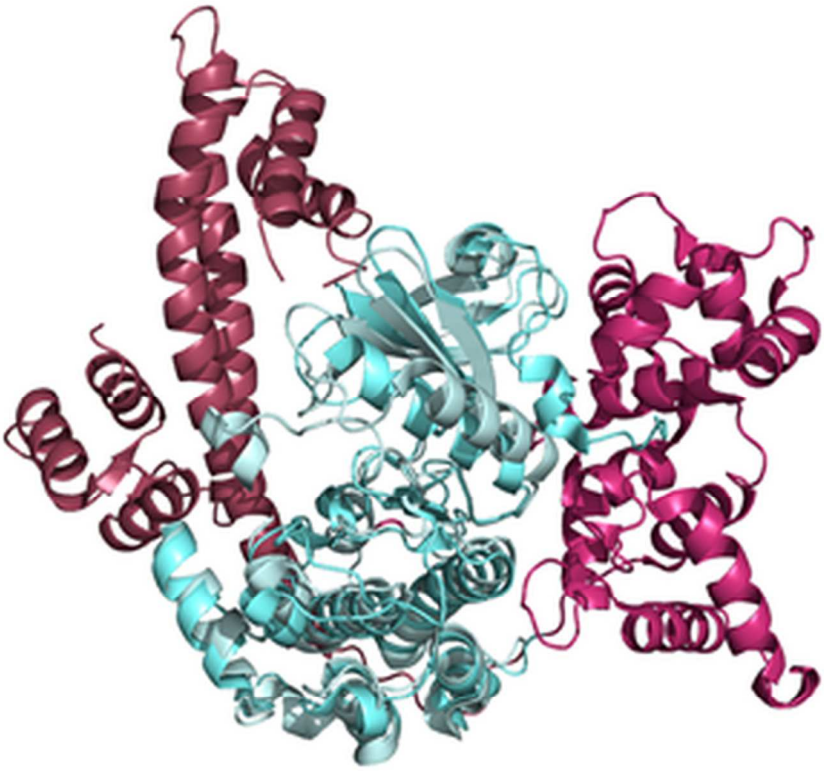


Figure 2: Ribbon diagram of the least-squares superposition of the inactive and active forms of CDPK1 with the kinase domains shown in different shades of cyan, the regulatory domain in shades of red, respectively. Only the kinase domain was used to calculate the transformation matrix which was then applied to the entire protein chain.

80x63mm (300 x 300 DPI)

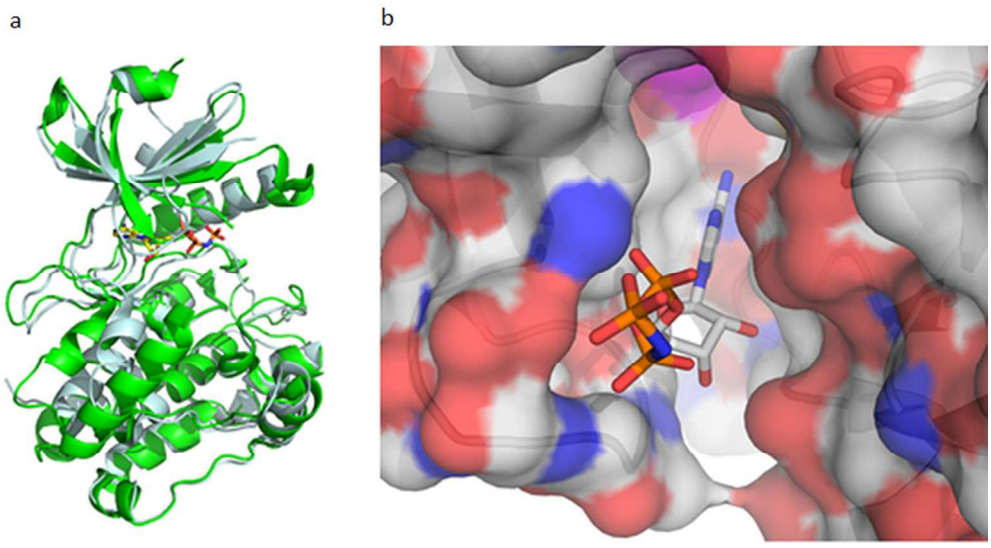


Figure 3: (a) Least squares superposition of the kinase domain of TgCDPK1 (depicted in cyan) in its active form (PDB: 3HX4) with HsCaMKII bound to an inhibitor (PDB: 2VZ6) (shown in green) (Rellos et al., 2010). The non-hydrolysable ATP analogue bound in CDPK1 is presented as ball-and-stick representation to highlight the ATP binding site. (b) Surface representation of TgCDPK1 viewing into the binding pocket with color coding according to atom type (oxygen in red, nitrogen in blue, carbon in grey). The surface of Gly128 (gatekeeper residue) is shown in magenta highlighting the additional space in the binding pocket of TgCDPK1.

67x36mm (300 x 300 DPI)

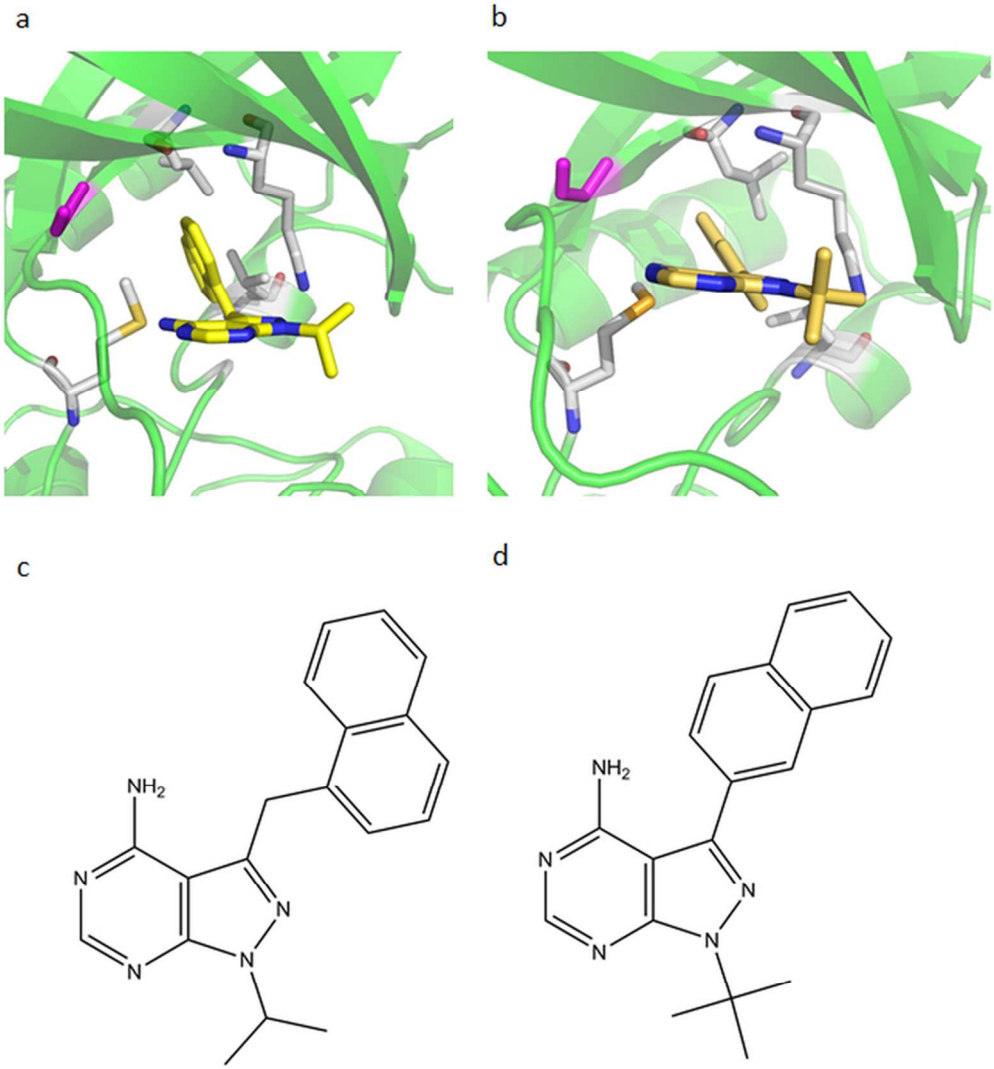


Figure 4: Close-up of BKIs bound to TgCDPK1 in the ATP binding site. The gatekeeper residue Gly128 is depicted in magenta, key hydrophobic residue of the binding site in grey (a) 1-(1-methylethyl)-3-(naphthalen-1-ylmethyl)-1H-pyrazolo[3,4-d]pyrimidin-4-amine shown in ball-and-stick representation (PDB: 3i7b) (b) 1-tert-butyl-3-naphthalen-2-yl-1H-pyrazolo[3,4-d]pyrimidin-4-amine (PDB:3i7c) (Ojo et al., 2010). (c) (d) chemical structures of the ligands.

80x86mm (300 x 300 DPI)

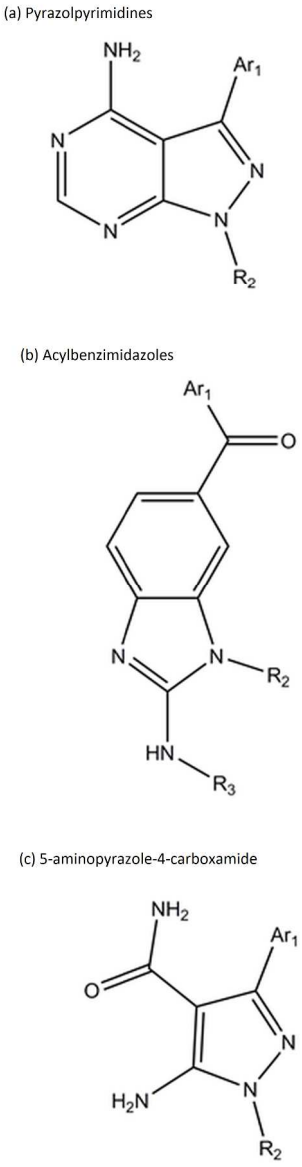


Figure 5: The three different scaffolds for TgCDPK1 inhibitors (a) Pyrazolpyrimidines (b) Acylbenzimidazoles (c) 5-aminopyrazole-4-carboxamide

110x372mm (300 x 300 DPI)

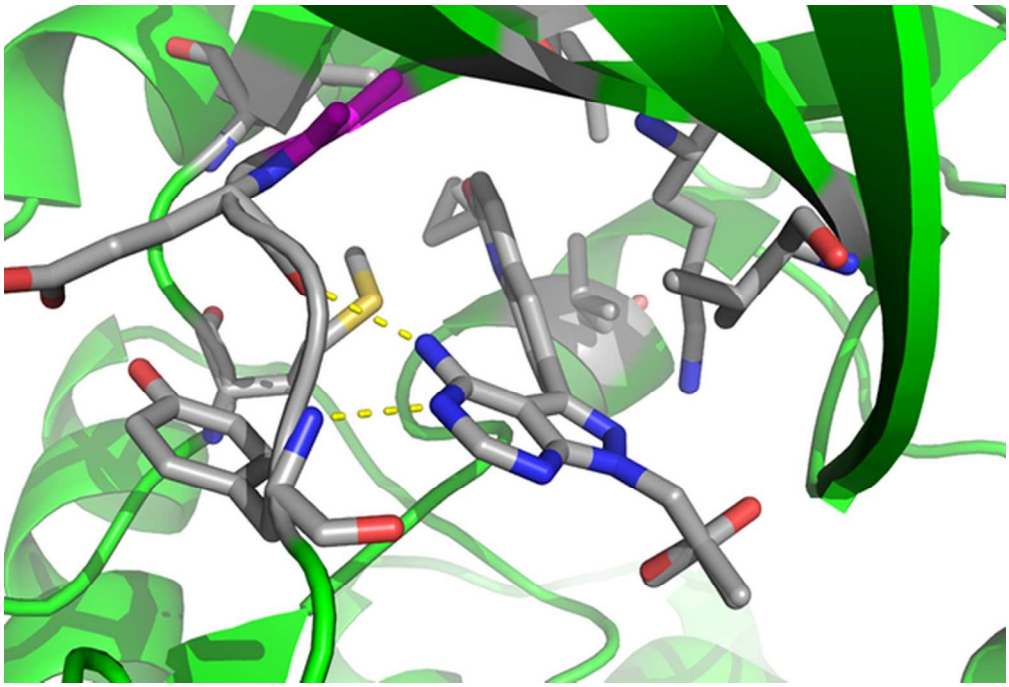


Figure 6: Crystal structure of (1-{4-amino-3-[2-(cyclopropyloxy)quinolin-6-yl]-1H-pyrazolo[3,4-d]pyrimidin-1-yl}-2-methylpropan-2-ol) shown in stick representation bound to for TgCDPK1 shown in cartoon representation with selected residues depicted in sticks (Vidadala et al., 2016).

80x54mm (300 x 300 DPI)

The Calcium-dependent protein kinase 1 from *Toxoplasma gondii* as target for structure-based drug design

Emily M. Cardew[1], Christophe L.M.J. Verlinde[2], Ehmke Pohl [1,3,4,*].

[1] Department of Biosciences, Durham University, Lower Mountjoy Durham DH1 3LE, UK

[2] Department of Biochemistry, University of Washington, Seattle, Washington, WA 98195, USA.

[3] Department of Chemistry, Durham University, South Road, Durham DH1, 3LE, UK.

[4] Biophysical Sciences Institute, Durham University, Durham DH1 3LE, UK.

* corresponding author, email: ehmke.pohl@durham.ac.uk

Abstract

The apicomplexan protozoan parasites include the causative agents of animal and human diseases ranging from malaria (Plasmodium spp.) to toxoplasmosis (*Toxoplasma gondii*). The complex life cycle of *T. gondii* is regulated by a unique family of calcium-dependent protein kinases (CDPKs) that have become the target of intensive efforts to develop new therapeutics. In this review, we will summarize structure-based strategies and recent successes in the pursuit of specific and selective inhibitors of *T. gondii* CDPK1.

Introduction

The phylum of Apicomplexa contains approximately 6000 unicellular, eukaryotic parasites including Plasmodium spp., the causative agent of Malaria, and *Toxoplasma gondii*, responsible for toxoplasmosis in many important farm animals and humans (Sato, 2011). Morphologically, all members of the apicomplexan family share a distinctive apical complex, together with species dependent apical-localised organelles. The apicoplast, for example, a non-photosynthetic plastid is found in both, Plasmodium spp. and *T. gondii* (McFadden & Yeh, 2017). These parasites employ complex life cycles including both sexual and asexual reproduction, and often involving multiple hosts. *T. gondii*, first described in 1908 and often regarded as one of the most successful apicomplexan parasites, represents the key model organism of the phylum (Weiss & Dubey, 2009, Szabo & Finney, 2017, Dubey, 2008). Its primary hosts are members of the Felidae (cats) family while all other warm-blooded animals, including humans, are intermediate hosts. It is estimated that up to one third of the human population is infected with *T. gondii* and thus are potential carriers. Although the infection is usually asymptotic in healthy individuals it can cause severe congenital disease during pregnancy (Kaye, 2011), and lead to life-threatening infections in immuno-compromised patients including those suffering from HIV, receiving an organ transplant or undergoing cancer chemotherapy treatment (Flegr *et al.*, 2014). Current toxoplasmosis treatment options are limited to a handful of antimicrobials such as sulphonamides, folic acid derivatives and certain macrolide antibiotics. However, these drugs often show limited efficacy and are associated with significant side effects (Alday & Doggett, 2017). Furthermore, there are no treatments available to target tissue cysts, the persistent form in which the parasite evades the host immune system, and to eradicate persistent *T. gondii* infections (Opsteegh *et al.*, 2015). Therefore, new drug targets and novel therapies are urgently needed. In addition to high-throughput screening approaches (Norcliffe *et al.*, 2014), structure-based methods in close combination with medicinal chemistry and biophysical and biological validation have become powerful tools in the search of new drugs and treatments (Hol, 2015, Groftehauge *et al.*, 2015, Muller, 2017, Verlinde *et al.*, 2009, Hunter, 2009).

The role of calcium-dependent protein kinases

In *T. gondii* Ca^{2+}-ions play key roles in cell signalling and in pathogen-host interactions including cell invasion, motility of the parasite within the host and differentiation during the parasites complex life cycle (Irvine, 1986, Nagamune *et al.*, 2008, Lourido & Moreno, 2015). Calcium dependent protein kinases (CDPKs) are a family of serine/threonine kinases that are only found in plants and protists including ciliates and apicomplexan parasites. Importantly, CDPKs provide the mechanistic link between calcium signalling and motility, differentiation and invasion (Tzen *et al.*, 2007, Billker *et al.*, 2009). These key roles of CDPKs have been proven through a range of knock-out studies in various species and underline the potential of CDPKs as targets for novel therapeutics (Long *et al.*, 2016). CDPKs are members of the Calmodulin/Calcium kinase (CaM) family. They share an N-terminal kinase domain (KD) linked via a junctional domain to a series of C-terminal Calcium-binding motifs.

In *T. gondii* at least twelve different CDPKs have been putatively identified ranging in size from 537 (CDPK3) to more than 2000 amino acids (CDPK7, CDPK80) (Morlon-Guyot *et al.*, 2014). The shared sequence identities range from 51% (CDPK1 and CDPK3) (Treeck *et al.*, 2014) to lower than 10% (Table 1) (Hui *et al.*, 2015). As evidenced by their vast variation in length and sequence, members of the CDPK family act upon dissimilar substrates and fulfil different functions in *T. gondii* biology. Recent knock-out studies using CRISPR-Cas9 indicate that CDPK4, CDPK5, CDPK6, CDPK8, and CDPK9, respectively, have no effect on virulence and on normal growth (Wang *et al.*, 2016), however, knock-down studies have shown that CDPK7 is crucial for survival due to a critical role in parasite division (Morlon-Guyot *et al.*, 2014). More detailed studies have been performed on the smaller family members. CDPK3 with 537 amino acids has been implicated in motility and host cell egress (McCoy *et al.*, 2017). CDPK2 (711 amino acids) has been shown to act as key regulator of amylopectin metabolism(Uboldi *et al.*, 2015). The loss of CDPK2 results in the build-up of amylum with fatal consequences for *T. gondii* in its chronic stage. Importantly, this family member contains an N-terminal carbohydrate-binding domain that may offer new opportunities for drug design (Uboldi *et al.*, 2015). CDPK1 (582 amino acids), which is mainly located in the cytosol, has been shown to be required for the microneme secretion at the apical complex and parasite proliferation. The molecular mechanism, however, remains elusive (Lourido *et al.*, 2010, Child *et al.*, 2017). Due to their key roles in infection CDPK1 from Plasmodium, Cryptosporidium and Toxoplasma spp. have attracted significant attention as a potential

novel drug target (Donald *et al.*, 2006, Sugi *et al.*, 2010, Larson *et al.*, 2012). Here we will review strategies and recent results in the discovery, design and potency of inhibitors targeting the kinase domain of CDPK1 from *T. gondii* (*Tg*CDPK1).

Activation of *Tg*CDPK1 by Calcium

The mechanism of activation and inhibition was unravelled in 2010 when the crystal structures of both the auto-inhibited and the Ca^{2+}-activated forms of *Tg*CDPK1 were published (Ojo *et al.*, 2010, Wernimont *et al.*, 2010). These structures revealed the expected kinase domain (KD) in similar overall conformations, however, the Ca^{2+}-binding domain (also designated CPDK activating domain or CAD) adopted two vastly different conformations and orientations (Figure 1). In its inactive state the CAD (shown in raspberry red) adopts an elongated form reminiscent of apo-calmodulin starting with a long helix followed by the first Ca^{2+}-binding motifs (EF-hands) which is connected via another long helix to the second pair of C-terminal EF-hands (Figure 1a). The first long helix has been suggested to be responsible for the auto-inhibitory effect by blocking the substrate binding site and providing a basic lysine residue to bind a cluster of conserved acidic residues. However, this may not be the only mechanism of deactivation as it has more recently been shown that removal of the regulatory domain alone does not lead to an active kinase domain (Ingram *et al.*, 2015). The CAD domain activated by Ca^{2+}-binding appears to be required to maintain the KD in its active conformation. Calcium binding leads to a dramatic rearrangement and refolding of the protein chain (Figure 1b) (Wernimont *et al.*, 2010). The entire regulatory domain is shifted to the other side of the protein hence liberating the active site of the kinase domain as shown in Figure 1c. In addition, the regulatory calcium-binding domain is collapsed so that the two long helices are no longer arranged in an anti-parallel fashion but are partially unwound and interwoven to form a more globular overall shape. These structural changes are reminiscent to the calcium-bound structure of calmodulin (Kursula, 2014).

Comparison with human kinases

Historically, characterising (protozoan) kinases as potential drug targets and developing selective inhibitors has been considered challenging due to the fact that the overall protein fold and the active sites are structurally well conserved (Scapin, 2002). The structural similarities are obvious when comparing the crystal structures of the kinase domain of

*Tg*CDPK1 with the Calcium/Calmodulin (CaM) dependent-kinase II from *H. sapiens* (*Hs*CaMKII) (Figure 2a) (Rellos *et al.*, 2010). These two proteins, which share a sequence identity of approximately 42% over 264 residues of the kinase domain, display the same canonical kinase fold and superimpose with an overall root mean square deviation of approximately 1.5 Å. Note that the loop over the adenosine triphosphate (ATP) binding site adopts a very different conformation presumably due to an induced fit of binding of two very different ligands. *Tg*CDPK1 is bound to the ATP analogue ANP (Figure 2a) while *Hs*CaMKII is bound to a comparatively small inhibitor. More importantly there are significant differences in the ATP binding site, specifically an unusually small residue (glycine) close to the adenine binding position. This residue, Gly128 is also termed the *gatekeeper* residue. Almost all mammalian kinases possess a large residue, a phenylalanine in *Hs*CaMKII for example, in this position. Hence, CDPK1 feature an enlarged ATP binding site with a hydrophobic pocket that can be exploited for structure-based drug design. This key structural difference in the binding pocket is shown in the surface representation where the ATP-analogue is shown as stick representation (Figure 2b). The additional space at the end of the pocket below the surface of the gatekeeper residue Gly 128 in magenta is clearly visible.

Development of specific *Tg*CDPK1 inhibitors

Soon after the structural differences were identified between *Tg*CDPK1 and the mammalian homologues two groups started to develop selective *Tg*CDPK1 inhibitors (Wernimont *et al.*, 2010, Ojo *et al.*, 2010). Initial compounds were based on known inhibitors previously developed for yeast kinases featuring amino acids with small side chains at the *gatekeeper* position. Importantly, these known kinase inhibitors, termed bumped kinase inhibitors (BKI) have been shown to be inactive against mammalian kinases (Hanke *et al.*, 1996). Generally, BKIs are based on the planar pyrazolo[3,4-d]pyrimidin-4-amine substituted with a bulky hydrophobic group on the C3 position (Bishop *et al.*, 1998). The first example of a BKI with a sub-µmolar IC_{50} is 1-(1-methylethyl)-3-(naphthalen-1-ylmethyl)-1H-pyrazolo[3,4-d] pyrimidin-4-amine. The co-crystal structure of TgCDPK1 shows that the naphtalen-1-ylmethyl- moiety fills the hydrophobic pocket created by the small gatekeeper residue Gly128 and lined by methionine and leucine residues, and one lysine residue (Figure 3a,b). The chemically closely related 1-tert-butyl-3-naphthalen-2-yl-1H-pyrazolo[3,4-d]pyrimidin-4-

amine (Figure 3c,d) adopts a similar conformation with the bulky aromatic substituent at the C3 position occupying the space next to the gatekeeper residue. Critically for the subsequent drug development was the fact that these and related BKIs reduced *T. gondii* proliferation significantly (Ojo *et al.*, 2010, Sugi *et al.*, 2010). These results sparked extensive medicinal chemistry efforts where a large number of compounds based on the BKI scaffold (4-amino-1H-pyrazole[3,4-d]pyrimidine) were synthesized and tested resulting in optimized *Tg*CDPK1 inhibitors. A number of compounds exhibited sub- or low-nanomolar IC_{50} values and high activity in parasite growth models (EC_{50} in the low- and sub-μmolar range) while retaining specificity when compared to mammalian kinases (Lourido *et al.*, 2013) (Zhang *et al.*, 2014) (Moine *et al.*, 2015). In addition to the pyrazolopyrimidine (PP) scaffolds, acylbenzimidazole and 5-aminopyrzazole-4-carboxamide based compounds have been shown to have similar properties (Figure 4) (Zhang *et al.*, 2012, Zhang *et al.*, 2014, Huang *et al.*, 2015). While the initial BKIs showed excellent potency *in vitro* and *in vivo* they also exhibited significant hERG (**h**uman **E**ther-**R**elated **G**ene) inhibition thus posing potential cardiotoxicity (Doggett *et al.*, 2014). Further extensive medicinal chemistry efforts finally led to the current lead *Tg*CDPK1 inhibitor, (1-{4-amino-3-[2-(cyclopropyloxy)quinolin-6-yl]-1H-pyrazolo[3,4-d]pyrimidin-1-yl}-2-methylpropan-2-ol) that combines high activity and selectivity with favourable pharmacokinetic properties and low hERG activity (Vidadala *et al.*, 2016). Note that the compound is bound to the protein via H-bonds of the pyrimidin ring to the main chain, while the hydrophobic cyclopropyloxy-quinoline moiety forms a large number of hydrophobic interactions (Figure 5). Taken together, the structure based approaches of drug development applied to *Tg*CDPK1 has led to three different series of compounds with high inhibitory activity, good pharmacokinetic parameters and promising efficacy in murine models.

Future challenges

Over the last five years there has been significant progress in the development of selective inhibitors of one of the key CDPKs from *T. gondii* achieved by taking advantage of a series of high-resolution crystal structures. While most of the previous work has focused on *T. gondii*, further work is currently underway to investigate inhibitors of CDPK1 from Cryptosporidium and Plasmodium spp. (Gaji *et al.*, 2014, Green *et al.*, 2015, Crowther *et al.*, 2016). Although the most promising *Tg*CDPK1 inhibitors show high efficacy in murine models more work

needs to be done to increase solubility and bio-availability in order to proceed to clinical trials. Furthermore, current lead compounds only target the ATP binding site of *Tg*CDPK1. However, allosteric kinase inhibitors and modulators have shown enormous potential to target specific kinases and could be further exploited (Fang *et al.*, 2013). Additional binding sites in less conserved regions such as the carbohydrate binding site recently discovered in *Tg*CDPK2 can serve as starting points for the development of new inhibitors (Uboldi *et al.*, 2015). Clearly, more works needs to be done to understand the role of the other members of the Apicomplexan CDPK family. In this regard, the recent development of CRISPR/Cas9 technology to modify the genes of members of the Apicomplexan family (Shen *et al.*, 2014, Vinayak *et al.*, 2015) will greatly facilitate the detailed analysis of the biological function of CDPK family members (Long *et al.*, 2016, Wang *et al.*, 2016).

Acknowledgments

This work was generously supported by the BBSRC grant BB/M024156/1. In addition, we would like to thank the Biophysical Sciences Institute (BSI) for seed corn funding and the Institute for Advanced Studies (IAS) for a Senior Research Fellowship at Durham University to CLMV. EMC is grateful for a studentship by the Newcastle-Liverpool-Durham BBSRC Doctoral Training Partnership.

Figure Legends

Figure 1: Ribbon representation of the crystal structure of CDPK1 from *T. gondii* with the kinase domain depicted in cyan, the regulatory domain in raspberry red (a) CDPK1 in its inactive auto-inhibited state (PDB code: 3KU2) (Wernimont *et al.*, 2010) (b) CDPK1 in its calcium-bound, activated state with the Ca^{2+}-ions shown as green spheres and the non-hydrolysable ligand ANP in stick representation (PDB code: 3HX4) (Wernimont *et al.*, 2010), (c) Ribbon diagram of the least-squares superposition of the inactive and active forms of *Tg*CDPK1 with the kinase domains shown in cyan (active) and blue (inactive), the regulatory domain in shades of red, respectively. Only the kinase domain was used to calculate the transformation matrix which was then applied to the entire protein chain.

Figure 2: (a) Least squares superposition of the kinase domain of *Tg*CDPK1 (depicted in cyan) in its active form with *Hs*CaMKII bound to an inhibitor (PDB: 2VZ6) (shown in orange) (Rellos *et al.*, 2010). The non-hydrolysable ATP analogue bound in *Tg*CDPK1 is presented as ball-and-stick representation to highlight the substrate binding site. (b) Surface representation of *Tg*CDPK1 viewing into the binding pocket with color coding according to atom type (oxygen in red, nitrogen in blue, carbon in grey). The surface of Gly128 (gatekeeper residue) is shown in magenta at the top of the figure highlighting the additional space in the binding pocket.

Figure 3: Close-up of BKIs bound to *Tg*CDPK1 in the ATP binding site. The gatekeeper residue Gly128 is depicted in magenta, key hydrophobic residue of the binding site are labelled and shown in grey (a) 1-(1-methylethyl)-3-(naphthalen-1-ylmethyl)-1H-pyrazolo[3,4-d]pyrimidin-4-amine shown in ball-and-stick representation (PDB: 3i7b) (b) chemical structure of the ligand (c) 1-tert-butyl-3-naphthalen-2-yl-1H-pyrazolo[3,4-d]pyrimidin-4-amine (PDB:3i7c) (Ojo *et al.*, 2010) (d) chemical structure of the ligand

Figure 4: The three different scaffolds for *Tg*CDPK1 inhibitors (a) Pyrazolpyrimidines (b) Acylbenzimidazoles (c) 5-aminopyrazole-4-carboxamide

Figure 5: Crystal structure of (1-{4-amino-3-[2-(cyclopropyloxy)quinolin-6-yl]-1H-pyrazolo[3,4-d]pyrimidin-1-yl}-2-methylpropan-2-ol) shown in stick representation bound to

for *Tg*CDPK1 shown in cartoon representation with selected residues depicted in sticks (Vidadala *et al.*, 2016).

References

Alday, P. H. & Doggett, J. S. (2017). *Drug Des Devel Ther* **11**, 273-293.

Billker, O., Lourido, S. & Sibley, L. D. (2009). *Cell Host Microbe* **5**, 612-622.

Bishop, A. C., Shah, K., Liu, Y., Witucki, L., Kung, C. & Shokat, K. M. (1998). *Curr Biol* **8**, 257-266.

Child, M. A., Garland, M., Foe, I., Madzelan, P., Treeck, M., van der Linden, W. A., Oresic Bender, K., Weerapana, E., Wilson, M. A., Boothroyd, J. C., Reese, M. L. & Bogyo, M. (2017). *MBio* **8**.

Crowther, G. J., Hillesland, H. K., Keyloun, K. R., Reid, M. C., Lafuente-Monasterio, M. J., Ghidelli-Disse, S., Leonard, S. E., He, P., Jones, J. C., Krahn, M. M., Mo, J. S., Dasari, K. S., Fox, A. M., Boesche, M., El Bakkouri, M., Rivas, K. L., Leroy, D., Hui, R., Drewes, G., Maly, D. J., Van Voorhis, W. C. & Ojo, K. K. (2016). *PLoS One* **11**, e0149996.

Doggett, J. S., Ojo, K. K., Fan, E., Maly, D. J. & Van Voorhis, W. C. (2014). *Antimicrob Agents Chemother* **58**, 3547-3549.

Donald, R. G., Zhong, T., Wiersma, H., Nare, B., Yao, D., Lee, A., Allocco, J. & Liberator, P. A. (2006). *Mol Biochem Parasitol* **149**, 86-98.

Dubey, J. P. (2008). *J Eukaryot Microbiol* **55**, 467-475.

Fang, Z., Grutter, C. & Rauh, D. (2013). *ACS Chem Biol* **8**, 58-70.

Flegr, J., Prandota, J., Sovickova, M. & Israili, Z. H. (2014). *PLoS One* **9**, e90203.

Gaji, R. Y., Checkley, L., Reese, M. L., Ferdig, M. T. & Arrizabalaga, G. (2014). *Antimicrob Agents Chemother* **58**, 2598-2607.

Green, J. L., Moon, R. W., Whalley, D., Bowyer, P. W., Wallace, C., Rochani, A., Nageshan, R. K., Howell, S. A., Grainger, M., Jones, H. M., Ansell, K. H., Chapman, T. M., Taylor, D. L., Osborne, S. A., Baker, D. A., Tatu, U. & Holder, A. A. (2015). *Antimicrob Agents Chemother* **60**, 1464-1475.

Groftehauge, M. K., Hajizadeh, N. R., Swann, M. J. & Pohl, E. (2015). *Acta Crystallogr D Biol Crystallogr* **71**, 36-44.

Hanke, J. H., Gardner, J. P., Dow, R. L., Changelian, P. S., Brissette, W. H., Weringer, E. J., Pollok, B. A. & Connelly, P. A. (1996). *J Biol Chem* **271**, 695-701.

Hol, W. G. (2015). *Acta Crystallogr F Struct Biol Commun* **71**, 485-499.

Huang, W., Ojo, K. K., Zhang, Z., Rivas, K., Vidadala, R. S., Scheele, S., DeRocher, A. E., Choi, R., Hulverson, M. A., Barrett, L. K., Bruzual, I., Siddaramaiah, L. K., Kerchner, K. M., Kurnick, M. D., Freiberg, G. M., Kempf, D., Hol, W. G., Merritt, E. A., Neckermann, G., de Hostos, E. L., Isoherranen, N., Maly, D. J., Parsons, M., Doggett, J. S., Van Voorhis, W. C. & Fan, E. (2015). *ACS Med Chem Lett* **6**, 1184-1189.

Hui, R., El Bakkouri, M. & Sibley, L. D. (2015). *Trends Pharmacol Sci* **36**, 452-460.

Hunter, W. N. (2009). *J Biol Chem* **284**, 11749-11753.

Ingram, J. R., Knockenhauer, K. E., Markus, B. M., Mandelbaum, J., Ramek, A., Shan, Y., Shaw, D. E., Schwartz, T. U., Ploegh, H. L. & Lourido, S. (2015). *Proc Natl Acad Sci U S A* **112**, E4975-4984.

Irvine, R. F. (1986). *Br Med Bull* **42**, 369-374.

Kaye, A. (2011). *J Pediatr Health Care* **25**, 355-364.

Kursula, P. (2014). *Amino Acids* **46**, 2295-2304.

Larson, E. T., Ojo, K. K., Murphy, R. C., Johnson, S. M., Zhang, Z., Kim, J. E., Leibly, D. J., Fox, A. M., Reid, M. C., Dale, E. J., Perera, B. G., Kim, J., Hewitt, S. N., Hol, W. G., Verlinde,

C. L., Fan, E., Van Voorhis, W. C., Maly, D. J. & Merritt, E. A. (2012). *J Med Chem* **55**, 2803-2810.
Long, S., Wang, Q. & Sibley, L. D. (2016). *Infect Immun* **84**, 1262-1273.
Lourido, S., Jeschke, G. R., Turk, B.E. & Sibley, D. (2013). *ACS Chemical Biology* **8**, 1155-1162.
Lourido, S. & Moreno, S. N. (2015). *Cell Calcium* **57**, 186-193.
Lourido, S., Shuman, J., Zhang, C., Shokat, K. M., Hui, R. & Sibley, L. D. (2010). *Nature* **465**, 359-362.
McCoy, J. M., Stewart, R. J., Uboldi, A. D., Li, D., Schroder, J., Scott, N. E., Papenfuss, A. T., Lehane, A. M., Foster, L. J. & Tonkin, C. J. (2017). *J Biol Chem* **292**, 7662-7674.
McFadden, G. I. & Yeh, E. (2017). *Int J Parasitol* **47**, 137-144.
Moine, E., Dimier-Poisson, I., Enguehard-Gueiffier, C., Loge, C., Penichon, M., Moire, N., Delehouze, C., Foll-Josselin, B., Ruchaud, S., Bach, S., Gueiffier, A., Debierre-Grockiego, F. & Denevault-Sabourin, C. (2015). *Eur J Med Chem* **105**, 80-105.
Morlon-Guyot, J., Berry, L., Chen, C. T., Gubbels, M. J., Lebrun, M. & Daher, W. (2014). *Cell Microbiol* **16**, 95-114.
Muller, I. (2017). *Acta Crystallogr D Struct Biol* **73**, 79-92.
Nagamune, K., Moreno, S. N., Chini, E. N. & Sibley, L. D. (2008). *Subcell Biochem* **47**, 70-81.
Norcliffe, J. L., Alvarez-Ruiz, E., Martin-Plaza, J. J., Steel, P. G. & Denny, P. W. (2014). *Parasitology* **141**, 8-16.
Ojo, K. K., Larson, E. T., Keyloun, K. R., Castaneda, L. J., Derocher, A. E., Inampudi, K. K., Kim, J. E., Arakaki, T. L., Murphy, R. C., Zhang, L., Napuli, A. J., Maly, D. J., Verlinde, C. L., Buckner, F. S., Parsons, M., Hol, W. G., Merritt, E. A. & Van Voorhis, W. C. (2010). *Nat Struct Mol Biol* **17**, 602-607.
Opsteegh, M., Kortbeek, T. M., Havelaar, A. H. & van der Giessen, J. W. (2015). *Clin Infect Dis* **60**, 101-107.
Rellos, P., Pike, A. C., Niesen, F. H., Salah, E., Lee, W. H., von Delft, F. & Knapp, S. (2010). *PLoS Biol* **8**, e1000426.
Sato, S. (2011). *Cell Mol Life Sci* **68**, 1285-1296.
Scapin, G. (2002). *Drug Discov Today* **7**, 601-611.
Shen, B., Brown, K. M., Lee, T. D. & Sibley, L. D. (2014). *MBio* **5**, e01114-01114.
Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J. D. & Higgins, D. G. (2011). *Mol Syst Biol* **7**, 539.
Sugi, T., Kato, K., Kobayashi, K., Watanabe, S., Kurokawa, H., Gong, H., Pandey, K., Takemae, H. & Akashi, H. (2010). *Eukaryot Cell* **9**, 667-670.
Szabo, E. K. & Finney, C. A. (2017). *Trends Parasitol* **33**, 113-127.
Treeck, M., Sanders, J. L., Gaji, R. Y., LaFavers, K. A., Child, M. A., Arrizabalaga, G., Elias, J. E. & Boothroyd, J. C. (2014). *PLoS Pathog* **10**, e1004197.
Tzen, M., Benarous, R., Dupouy-Camet, J. & Roisin, M. P. (2007). *Parasite* **14**, 141-147.
Uboldi, A. D., McCoy, J. M., Blume, M., Gerlic, M., Ferguson, D. J., Dagley, L. F., Beahan, C. T., Stapleton, D. I., Gooley, P. R., Bacic, A., Masters, S. L., Webb, A. I., McConville, M. J. & Tonkin, C. J. (2015). *Cell Host Microbe* **18**, 670-681.
Verlinde, C. L., Fan, E., Shibata, S., Zhang, Z., Sun, Z., Deng, W., Ross, J., Kim, J., Xiao, L., Arakaki, T. L., Bosch, J., Caruthers, J. M., Larson, E. T., Letrong, I., Napuli, A., Kelly, A., Mueller, N., Zucker, F., Van Voorhis, W. C., Buckner, F. S., Merritt, E. A. & Hol, W. G. (2009). *Curr Top Med Chem* **9**, 1678-1687.

Vidadala, R. S., Rivas, K. L., Ojo, K. K., Hulverson, M. A., Zambriski, J. A., Bruzual, I., Schultz, T. L., Huang, W., Zhang, Z., Scheele, S., DeRocher, A. E., Choi, R., Barrett, L. K., Siddaramaiah, L. K., Hol, W. G., Fan, E., Merritt, E. A., Parsons, M., Freiberg, G., Marsh, K., Kempf, D. J., Carruthers, V. B., Isoherranen, N., Doggett, J. S., Van Voorhis, W. C. & Maly, D. J. (2016). *J Med Chem* **59**, 6531-6546.

Vinayak, S., Pawlowic, M. C., Sateriale, A., Brooks, C. F., Studstill, C. J., Bar-Peled, Y., Cipriano, M. J. & Striepen, B. (2015). *Nature* **523**, 477-480.

Wang, J. L., Huang, S. Y., Li, T. T., Chen, K., Ning, H. R. & Zhu, X. Q. (2016). *Parasitol Res* **115**, 697-702.

Weiss, L. M. & Dubey, J. P. (2009). *Int J Parasitol* **39**, 895-901.

Wernimont, A. K., Artz, J. D., Finerty, P., Jr., Lin, Y. H., Amani, M., Allali-Hassani, A., Senisterra, G., Vedadi, M., Tempel, W., Mackenzie, F., Chau, I., Lourido, S., Sibley, L. D. & Hui, R. (2010). *Nat Struct Mol Biol* **17**, 596-601.

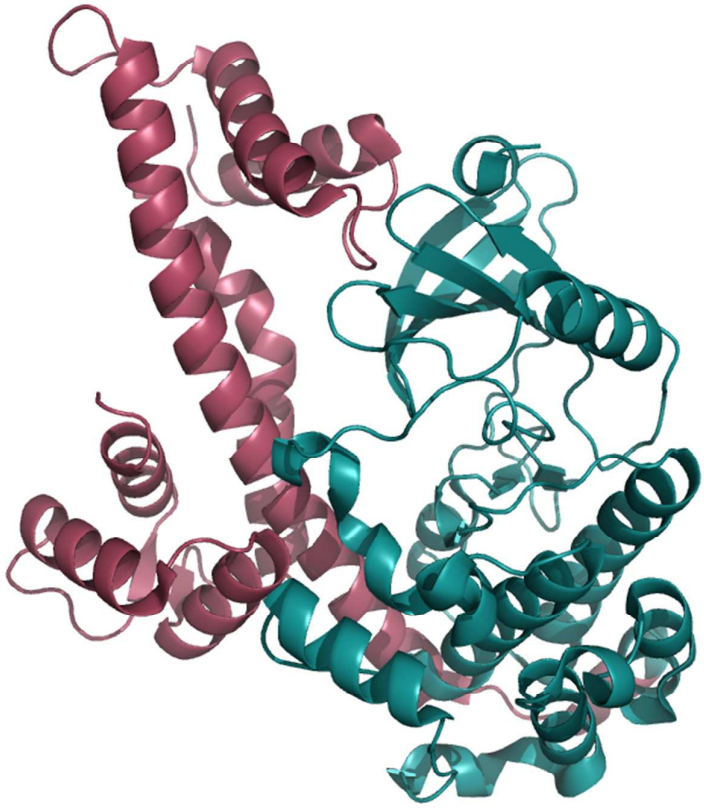
Zhang, Z., Ojo, K. K., Johnson, S. M., Larson, E. T., He, P., Geiger, J. A., Castellanos-Gonzalez, A., White, A. C., Jr., Parsons, M., Merritt, E. A., Maly, D. J., Verlinde, C. L., Van Voorhis, W. C. & Fan, E. (2012). *Bioorg Med Chem Lett* **22**, 5264-5267.

Zhang, Z., Ojo, K. K., Vidadala, R., Huang, W., Geiger, J. A., Scheele, S., Choi, R., Reid, M. C., Keyloun, K. R., Rivas, K., Siddaramaiah, L. K., Comess, K. M., Robinson, K. P., Merta, P. J., Kifle, L., Hol, W. G., Parsons, M., Merritt, E. A., Maly, D. J., Verlinde, C. L., Van Voorhis, W. C. & Fan, E. (2014). *ACS Med Chem Lett* **5**, 40-44.

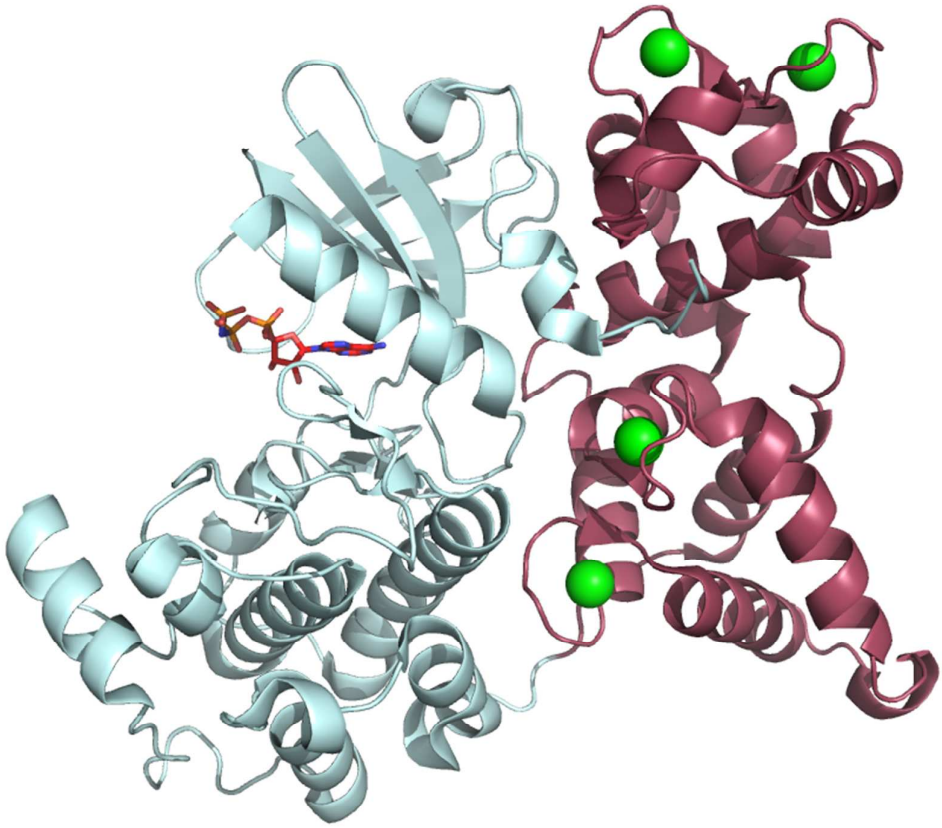
Table 1: The protein sequence identities between the 12 putative CDPKs of *T. gondii*, rounded to the nearest whole number, derived from a multiple sequence alignment (MSA) generated using Clustal Omega (Sievers *et al.*, 2011).

	CDPK1	CDPK2	CDPK2A	CDPK2B	CDPK3	CDPK4	CDPK4A	CDPK5	CDPK6	CDPK7	CDPK8	CDPK9
CDPK1		22%	23%	30%	51%	14%	24%	25%	9%	7%	6%	22%
CDPK2	22%		34%	34%	25%	14%	21%	32%	14%	7%	9%	26%
CDPK2A	23%	34%		37%	27%	20%	20%	32%	15%	9%	9%	24%
CDPK2B	30%	34%	37%		33%	16%	25%	36%	13%	8%	8%	26%
CDPK3	51%	25%	27%	33%		14%	25%	30%	12%	7%	7%	25%
CDPK4	14%	14%	20%	16%	14%		14%	15%	17%	11%	9%	14%
CDPK4A	24%	21%	20%	25%	25%	14%		21%	8%	6%	6%	18%
CDPK5	25%	32%	32%	36%	30%	15%	21%		13%	7%	8%	26%
CDPK6	9%	14%	15%	13%	12%	17%	8%	13%		16%	8%	12%
CDPK7	7%	7%	9%	8%	7%	11%	6%	7%	16%		8%	7%
CDPK8	6%	9%	9%	8%	7%	9%	6%	8%	8%	8%		7%
CDPK9	22%	26%	24%	26%	25%	14%	18%	26%	12%	7%	7%	

Figure 1
(a)



(b)



(c)

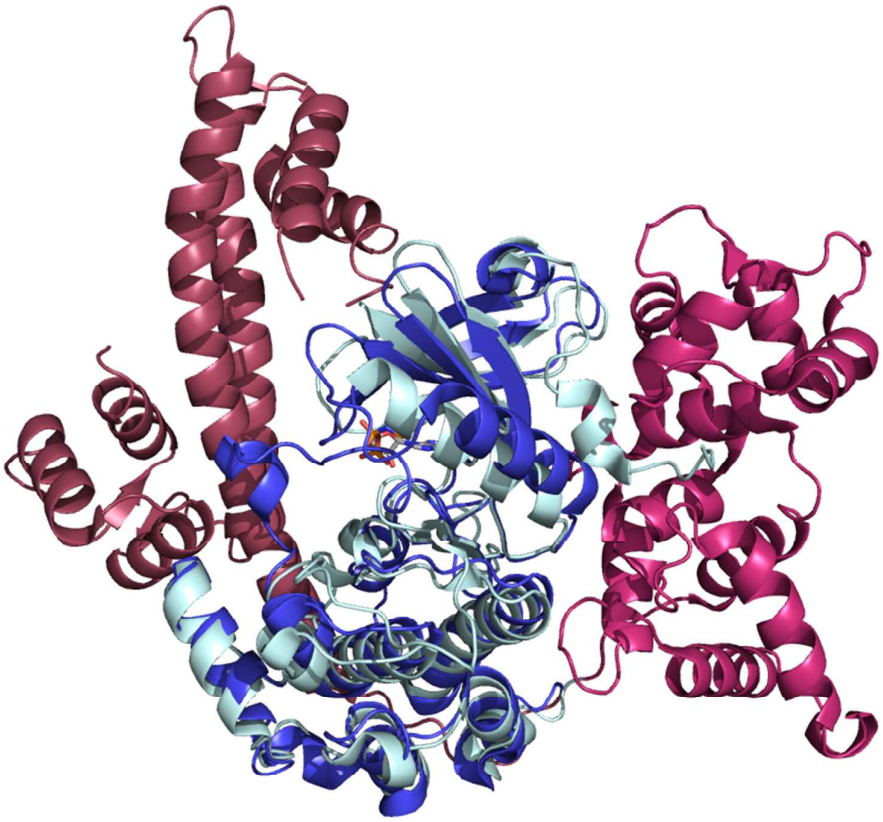


Figure 2

(a) (b)

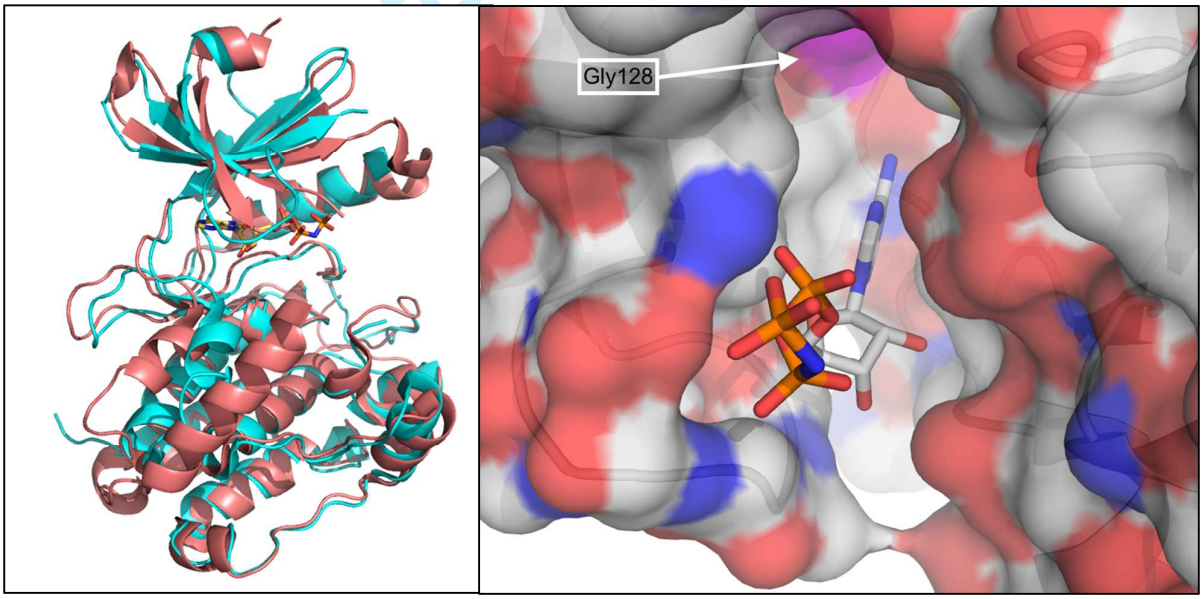
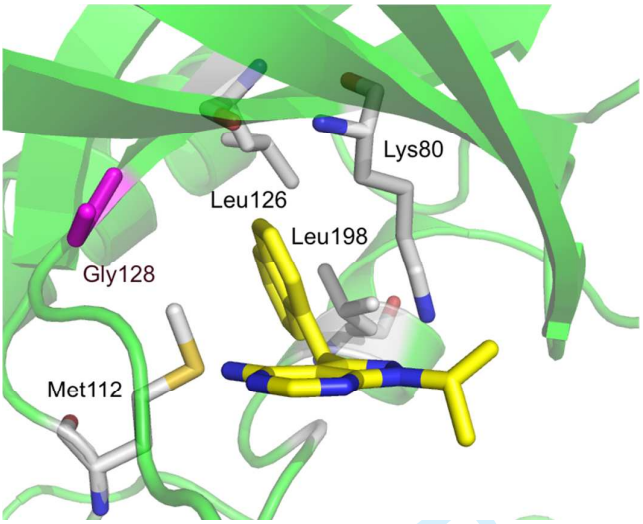
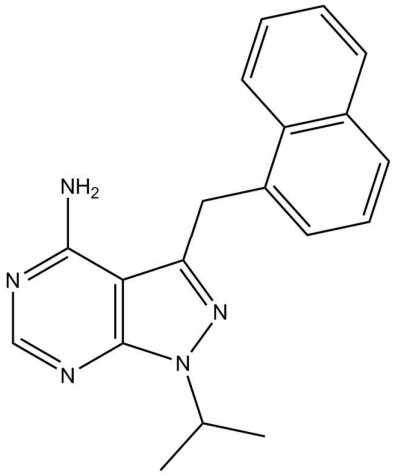


Figure 3

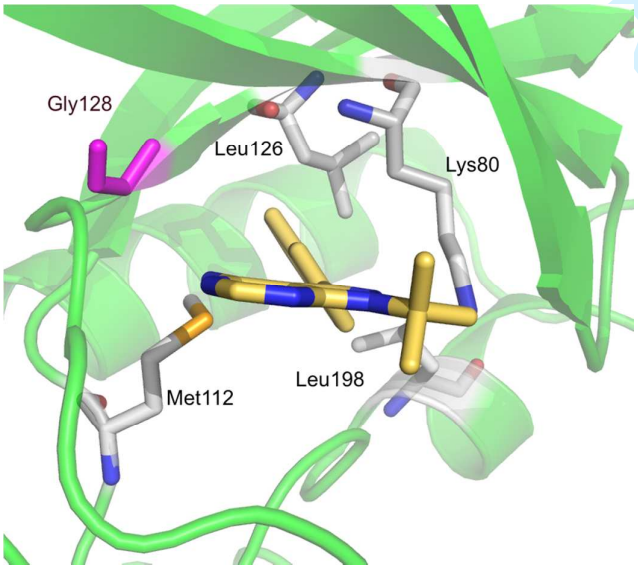
(a)



(b)



(c)



(d)

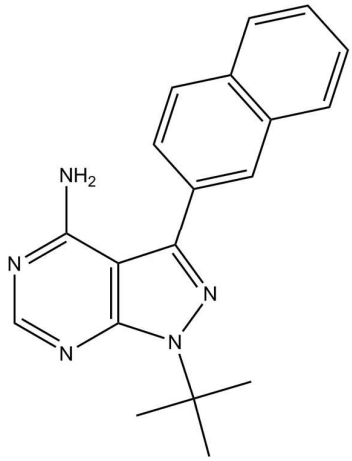
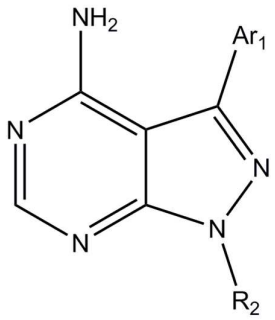
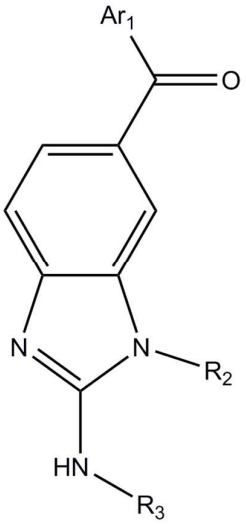


Figure 4

(a) Pyrazolpyrimidines



(b) Acylbenzimidazoles



(c) 5-aminopyrazole-4-carboxamide

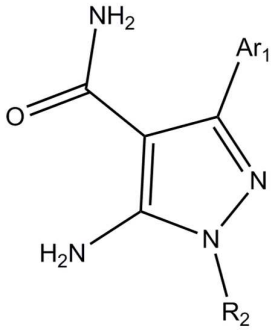
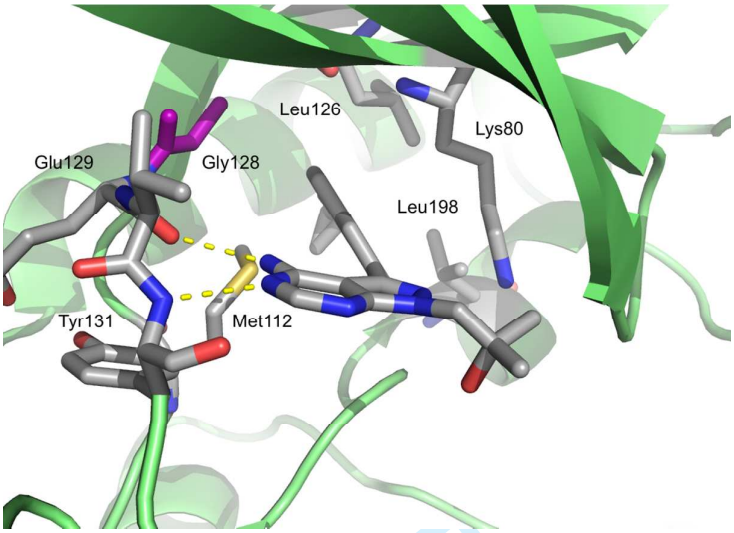
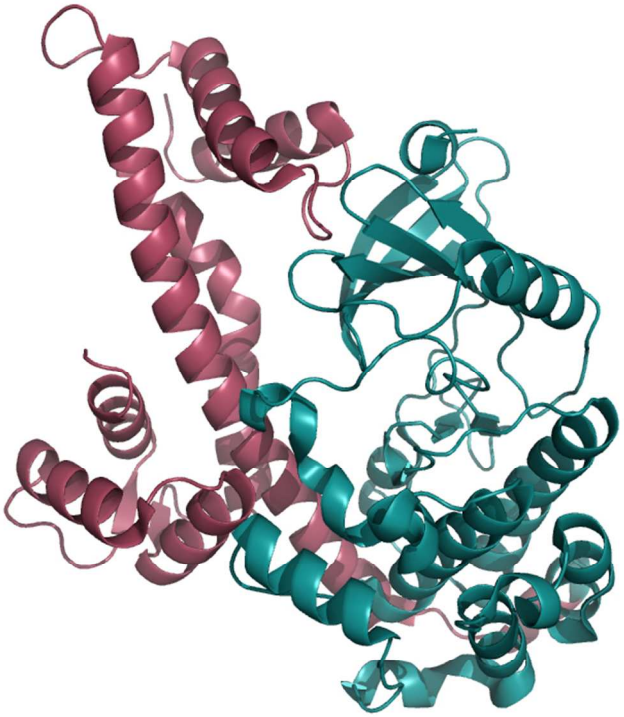


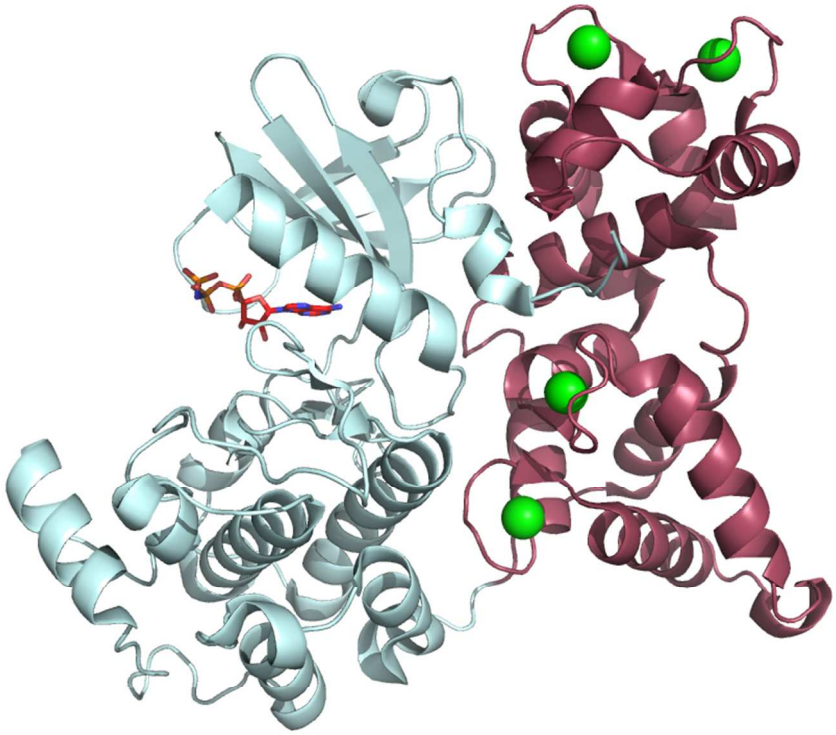
Figure 5.



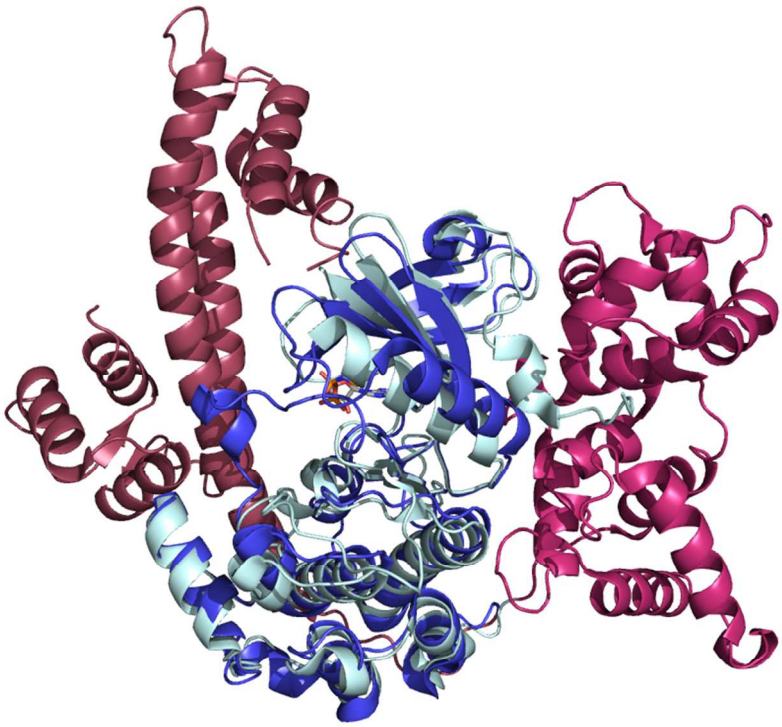


Ribbon representation of the crystal structure of CDPK1 from T. gondii with the kinase domain depicted in cyan, the regulatory domain in raspberry red (a) CDPK1 in its inactive auto-inhibited state (PDB code: 3KU2) (Wernimont et al., 2010) (b) CDPK1 in its calcium-bound, activated state with the Ca2+-ions shown as green spheres and the non-hydrolysable ligand ANP in stick representation (PDB code: 3HX4) (Wernimont et al., 2010), (c) Ribbon diagram of the least-squares superposition of the inactive and active forms of TgCDPK1 with the kinase domains shown in cyan (active) and blue (inactive), the regulatory domain in shades of red, respectively. Only the kinase domain was used to calculate the transformation matrix which was then applied to the entire protein chain.

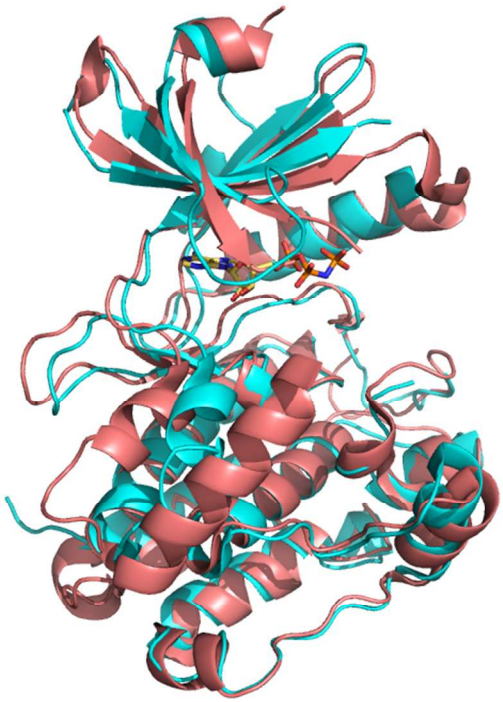
342x262mm (72 x 72 DPI)



339x271mm (72 x 72 DPI)

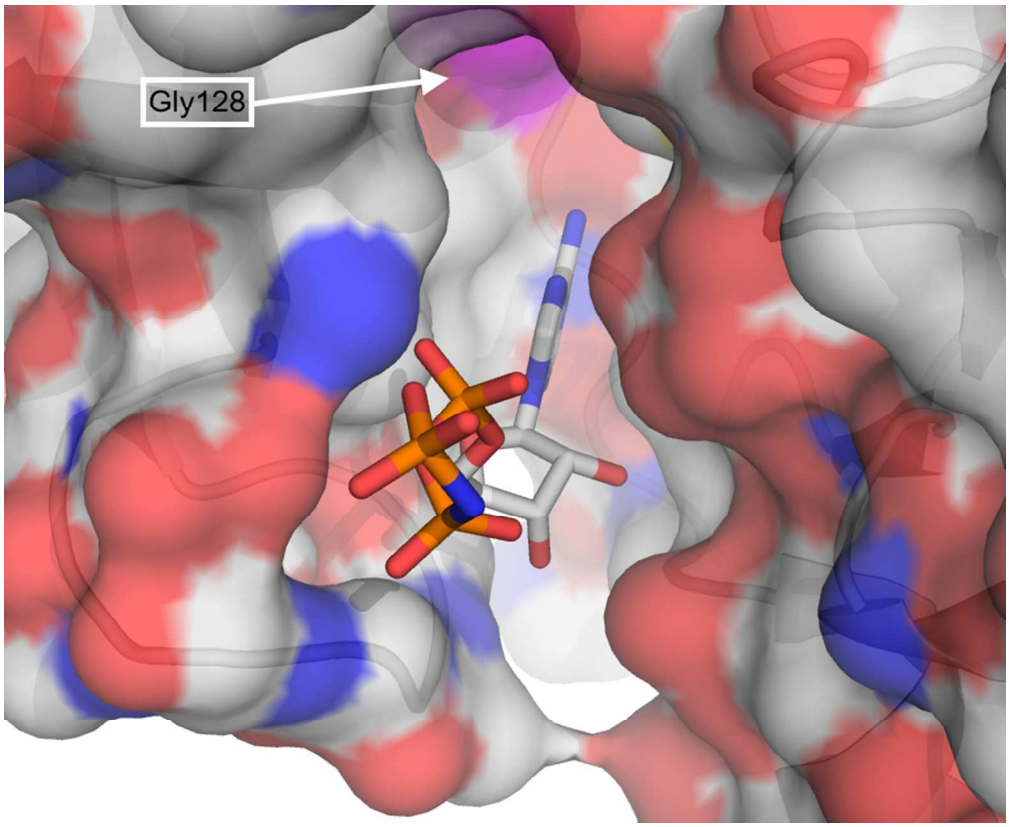


358x271mm (72 x 72 DPI)

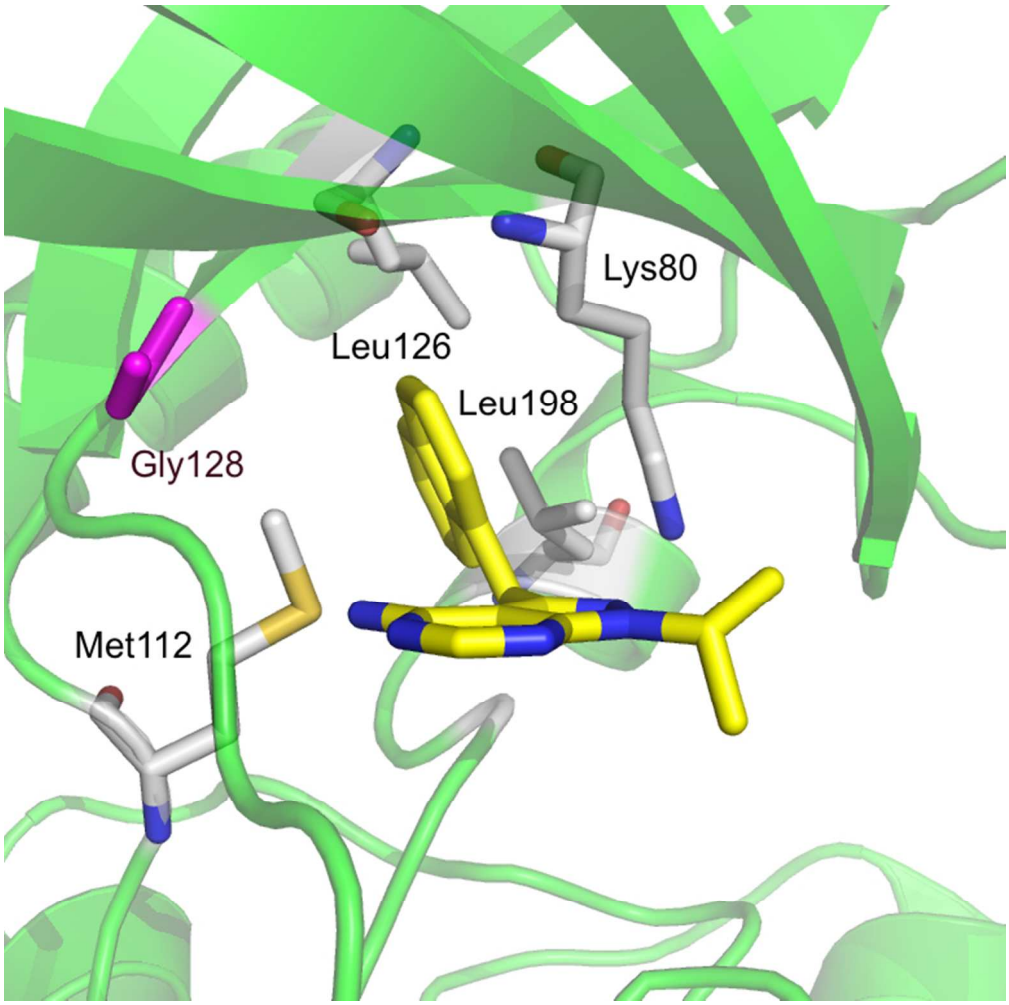


: (a) Least squares superposition of the kinase domain of TgCDPK1 (depicted in cyan) in its active form with HsCaMKII bound to an inhibitor (PDB: 2VZ6) (shown in orange) (Rellos et al., 2010). The non-hydrolysable ATP analogue bound in TgCDPK1 is presented as ball-and-stick representation to highlight the substrate binding site. (b) Surface representation of TgCDPK1 viewing into the binding pocket with color coding according to atom type (oxygen in red, nitrogen in blue, carbon in grey). The surface of Gly128 (gatekeeper residue) is shown in magenta at the top of the figure highlighting the additional space in the binding pocket.

330x268mm (72 x 72 DPI)

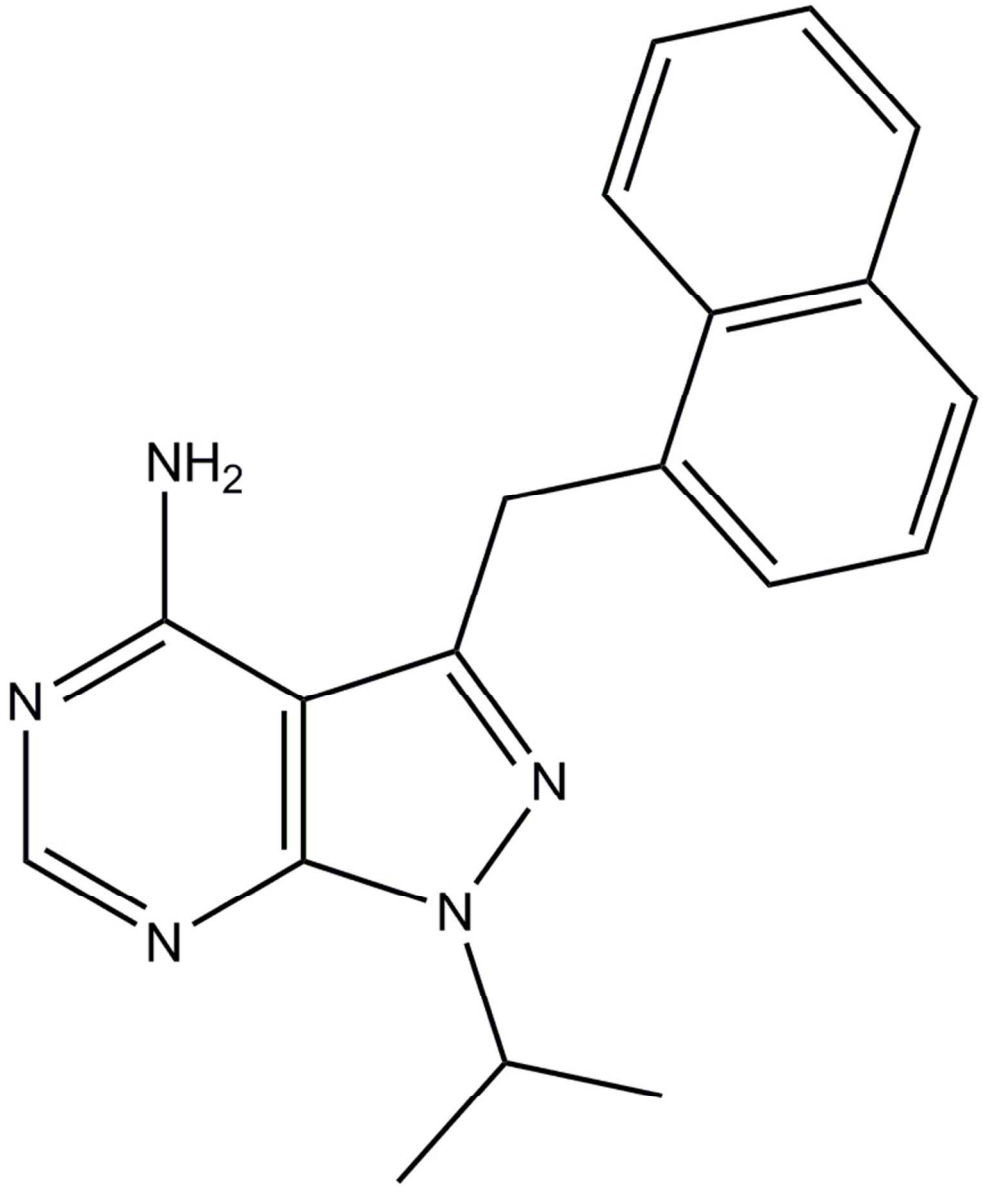


319x261mm (72 x 72 DPI)

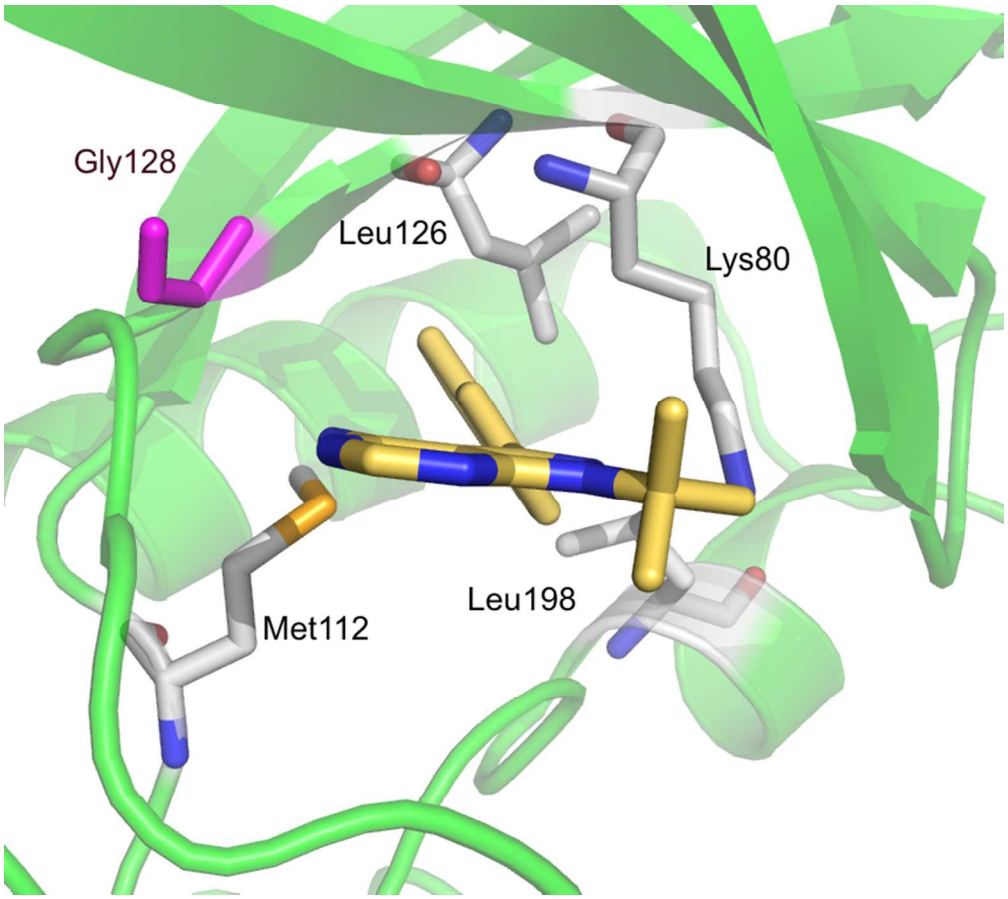


Close-up of BKIs bound to TgCDPK1 in the ATP binding site. The gatekeeper residue Gly128 is depicted in magenta, key hydrophobic residue of the binding site are labelled and shown in grey (a) 1-(1-methylethyl)-3-(naphthalen-1-ylmethyl)-1H-pyrazolo[3,4-d]pyrimidin-4-amine shown in ball-and-stick representation (PDB: 3i7b) (b) chemical structure of the ligand (c) 1-tert-butyl-3-naphthalen-2-yl-1H-pyrazolo[3,4-d]pyrimidin-4-amine (PDB:3i7c) (Ojo et al., 2010) (d) chemical structure of the ligand

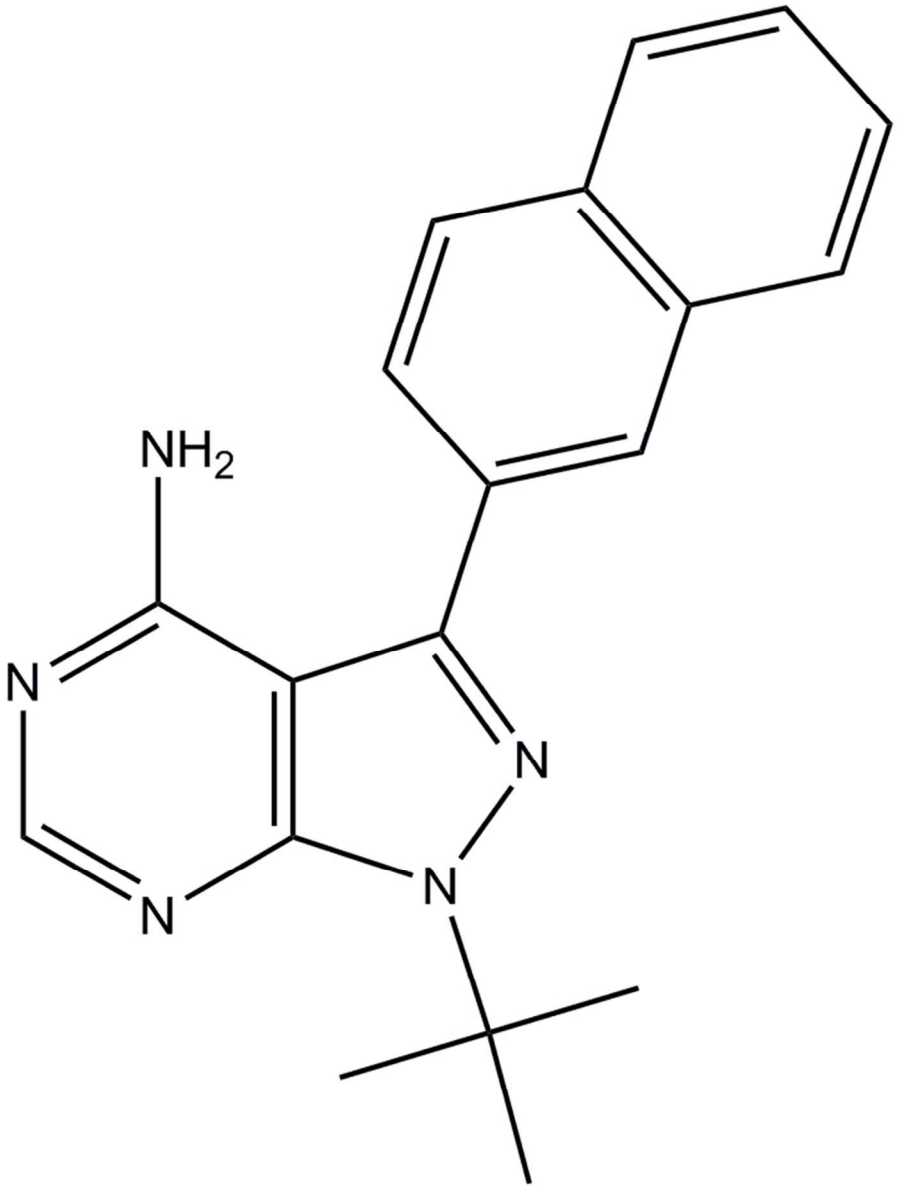
272x270mm (72 x 72 DPI)



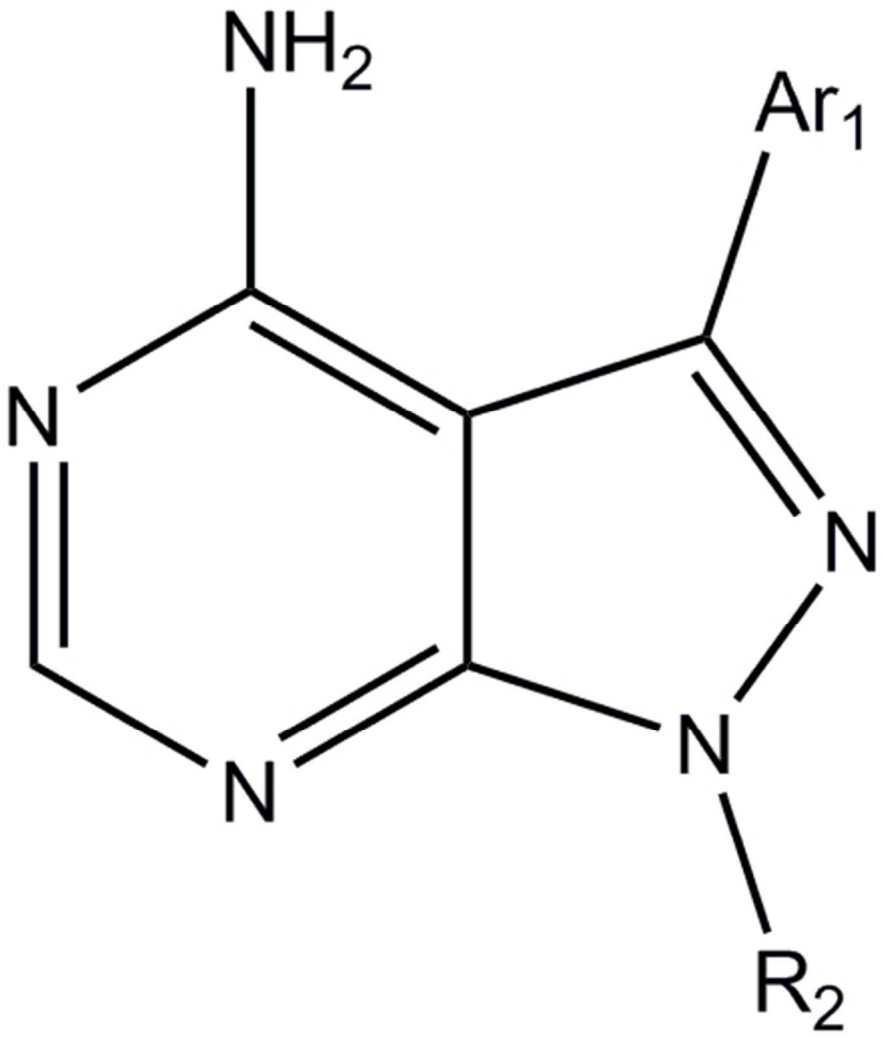
65x79mm (300 x 300 DPI)



301x267mm (72 x 72 DPI)

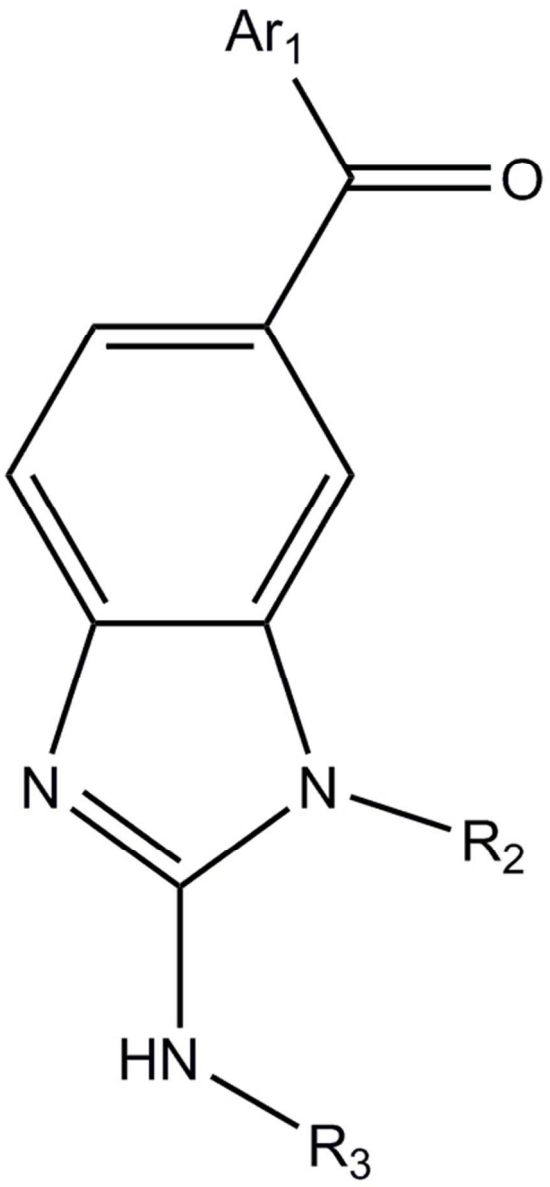


61x81mm (300 x 300 DPI)

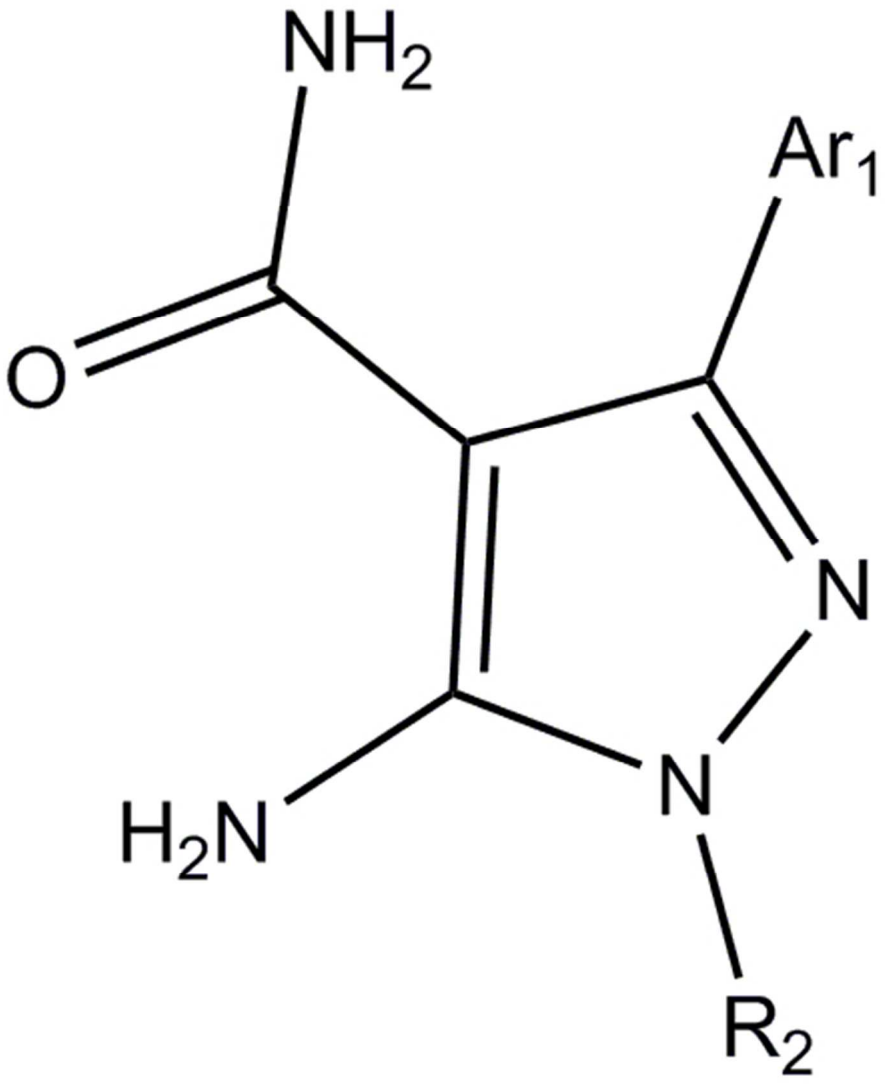


The three different scaffolds for TgCDPK1 inhibitors (a) Pyrazolpyrimidines (b) Acylbenzimidazoles (c) 5-aminopyrazole-4-carboxamide

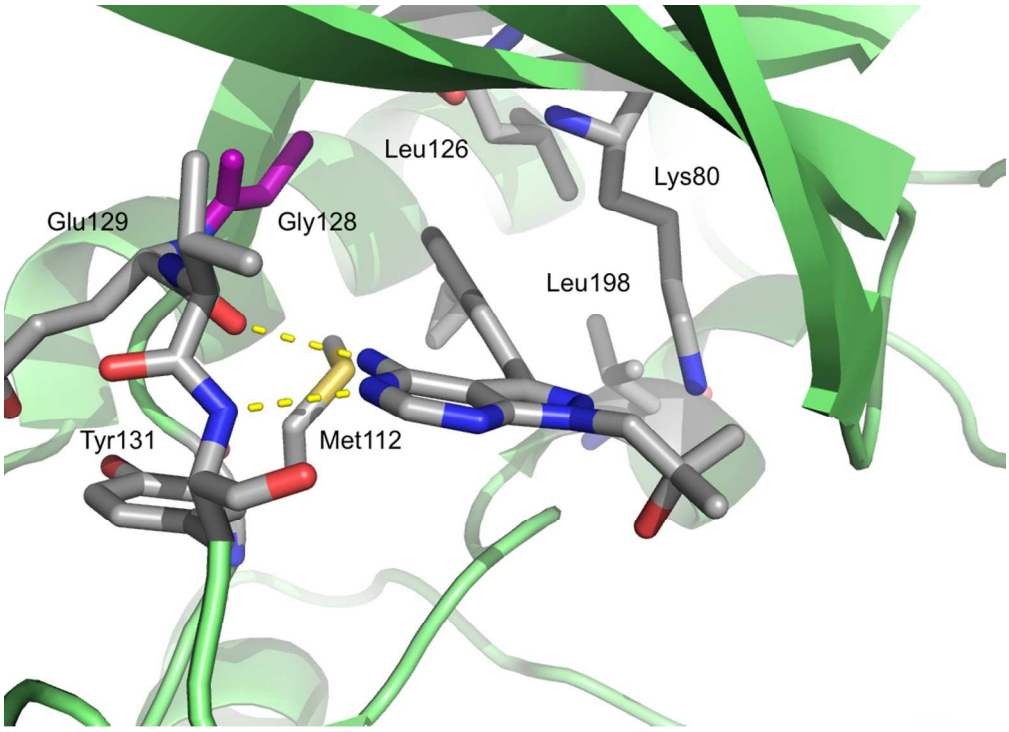
38x44mm (300 x 300 DPI)



34x74mm (300 x 300 DPI)



38x46mm (300 x 300 DPI)



Crystal structure of (1-{4-amino-3-[2-(cyclopropyloxy)quinolin-6-yl]-1H-pyrazolo[3,4-d]pyrimidin-1-yl}-2-methylpropan-2-ol) shown in stick representation bound to for TgCDPK1 shown in cartoon representation with selected residues depicted in sticks (Vidadala et al., 2016).

390x281mm (72 x 72 DPI)